On the Complex Non-Minimum Phase Zeros in Flexure Mechanisms

by

Leqing Cui

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in the University of Michigan
2017

Doctoral Committee:

Associate Professor Shorya Awtar, Chair
Professor James S. Freudenberg
Associate Professor Chinedum E. Okwudire
Professor Kon-Well Wang

Leqing Cui

Leqing@umich.edu

ORCID ID: 0000-0002-8779-4944

To my parents and my wife

ACKNOWLEDGEMENTS

I would like to thank my advisor, Prof. Shorya Awtar, for his guidance and support throughout my graduate studies. He is incredibly patient in teaching me everything, from the daily gramma to the complex academic research methodology. In all, he is my *shifu*.

I am grateful to Prof. James Freudenberg, Prof. Chinedum Okwudire, and Prof. Kon-Well Wang, for their willingness to serve in my doctoral committee and offering much needed suggestions, help, comments, and criticism for my research.

I would like to thank my colleagues at Precision System Design Lab for all their help and support. Part of the modeling work in Chapter 2 is done in collaboration with Dr. Gaurav Parmar. Dhanush Mariappan helped me with the manufacturing of the experimental setup. Kai Wu helped me with the system identification of the XY flexure stage. Marijn Nijenhuis helped me with the damping simulation of the flexure mechanisms. Dr. Mohammad Olfatina guided me into the world of MEMS fabrication and test. I would also like to thank Prof. Hongzhe Zhao, Ramiro Pinon, Adrian Adames, and David Hiemstra for their help.

I am also grateful to Toby Donajkowski, who offered me all the help in the Mechatronics Lab. I would like to thank Elizabeth Throckmorton, who patiently provided proof-reading for me and helped me improve my writing. Additionally, I would like to thank the M&M seminar held by three labs (directed by Prof. Shorya Awtar, Prof. Kira Barton, and Prof. Chinedum Okwudire), where I could learn from their expertise and develop my presentation skills.

Finally, I would like to thank my family, including my parents and my wife Yang, for their unconditional and endless support. Thanks for keeping me happy every day and offering me a harbor to return.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The objective of this thesis is to investigate the genesis and consequences of complex non-minimum phase (CNMP) zeros observed experimentally in the dynamics of large displacement flexure mechanisms. The knowledge gained through this research may be used to overcome traditional tradeoffs between large displacement and good dynamic performance in various applications that employ flexure mechanisms. Applications include high precision positioning stages for wafer inspection, motion stages for atomic force microscopy, and micro electro-mechanical systems for high speed scanning.

In the first part of this study, a lumped parameter approach is proposed to model the CNMP behavior. It is shown that CNMP zeros appear due to the inherent geometric nonlinearities in the mechanics of a simple beam (i.e., arc length conservation) as well as due to small structural asymmetry (e.g., manufacturing tolerances) at certain operating points in the mechanism's workspace. In addition, a hypothesis of a correlation between CNMP zeros and a phenomenon known as curve veering is proposed. This hypothesis is verified in later part of the thesis.

Second, to understand the physics of CNMP zeros, a systematic investigation is presented, assuming a broader scope of lightly damped flexible systems. The analysis of modal residues based on modal theory is conducted to find the conditions for the existence of each possible type of zeros (i.e., marginally minimum phase, real non-minimum phase, and complex non-minimum phase). To study the transition between these types of zeros, a graphical technique is proposed that utilizes the sequences of the dominant modes and the sequences of each decomposed mode's sign. It is shown that CNMP zeros appear in a system that is decomposed as two closely spaced modes and a remainder mode with alternating signs, and these modes dominate at their corresponding frequency regions. Furthermore, the impact of CNMP zeros on closed loop control, as well as the initial properties (e.g., initial undershoot) of the step response for a system with CNMP zeros are visualized via this technique.

Third, the correlation between CNMP zeros and curve veering is examined. It is shown under certain conditions that CNMP zeros appear in the presence of curve veering. The physical insights of these conditions are presented. This finding reveals a conditional connection between the two fields (i.e., CNMP and curve veering) that have been separately studied in the literature.

Next, an experimental setup is designed to validate the above findings. To conduct the validation, a design procedure is presented to minimize the extraneous factors. The frequency responses are then measured and match well with the model prediction.

Lastly, a mechatronic system (including mechanical system and control system) design approach is presented based on the knowledge developed above, in order to achieve better closed loop performance in large motion range flexure stages compared to the state-of-the-art stages. The study is based on a simplified system considered in this thesis. It is shown that a system with a sequence of alternating signs is easier to stabilize compared to a system with non-alternating signs, even when the former one has CNMP zeros. Therefore, it is recommended to design the mechanical system by studying the sequence of decomposed modes' signs (and the sequence of dominant modes) to achieve the best possible dynamic characteristics for good control performance. An example is presented with both simulation and experiment.

Chapter 1. Introduction and Background

1.1. Flexure for High Speed Motion Applications

Flexure mechanisms (Fig 1-1a) employ elastic deformation instead of rolling or sliding joints to provide guided motion along certain compliant directions and load bearing capability along the remaining stiff directions [1-6]. The term mechanism, in this context, signifies a similarity with traditional rigid-link mechanisms (Fig 1-1b), which also selectively allow certain motions and constrain others. For example, both the parallelogram mechanisms of Fig 1-1a and Fig 1-1b offer a translational degree of freedom (DoF) in the transverse direction and degrees of constraint (DoC) in the axial and rotational directions [3-4]. From another perspective, flexure mechanisms may also be considered similar to elastic structures (Fig 1-1c) since both represent an elastic continuum. But the key distinction between these two lies in the fact that, unlike flexure mechanisms, structures are typically designed for high stiffness and minimal motion along all directions [7].

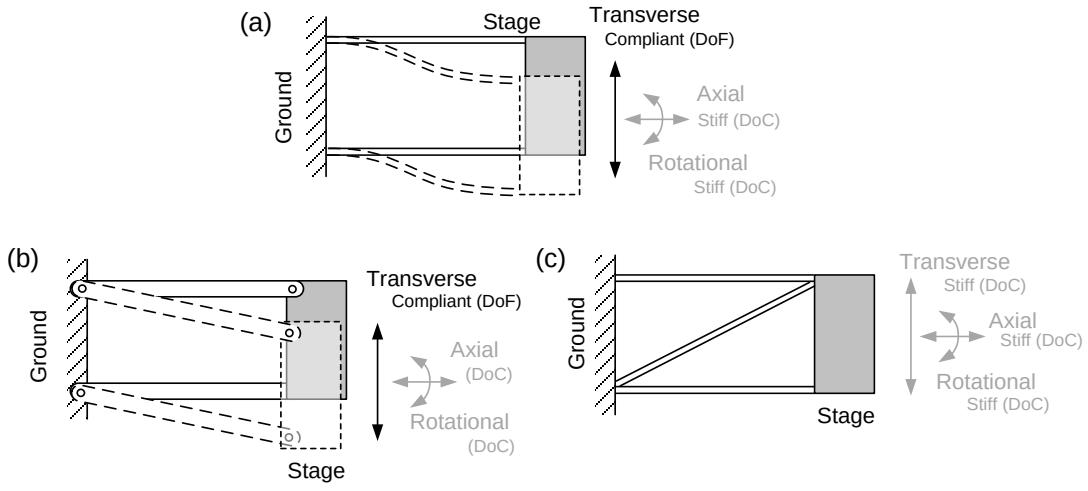

Fig 1-1 A comparison between different mechanisms: (a) parallelogram flexure mechanism; (b) parallelogram rigid mechanism; (c) elastic structure

This research investigation is motivated by the need to achieve large range, high precision, and high speed all simultaneously in multi-axis flexure-based motion systems [3, 8]. Such capability is of practical importance in various applications such as compact and affordable motion stages for semiconductor wafer inspection [9] and micro electro-mechanical systems (MEMS) scanners for high speed imaging [10]. Flexure mechanisms are well-suited for these applications because of their joint-less construction and inherently high precision due to lack of friction and backlash, but present significant tradeoffs between large displacement and dynamic performance [11].

1.2. Performance Tradeoff and Limitation

Large displacement generally implies transverse deformation (see Fig 1-1a) of the constituent beams in the flexure mechanism greater than 5% of the beam length. This typically corresponds to several millimeters of displacement or motion range for a desktop size flexure-based motion system. The relevant system dynamics include natural frequencies, mode shapes, and transfer functions. The closed-loop dynamic performance objectives include high bandwidth, good noise and disturbance rejection, good command tracking, small steady-state error, fast point-to-point positioning and settling, stability robustness, low sensitivity to plant variations, etc.

While recent results have demonstrated large range as well as high precision in multi-axis flexure mechanisms, achieving dynamic performance remains a challenge [11, 12]. Fig 1-2a shows an XY nano-positioning system based on a parallel-kinematic flexure mechanism, designed to achieve a range of 10 mm and a precision of ±25 nm per axis. This mechanism employs a systematic and symmetric layout of eight double parallelogram flexure modules (DPFM). This design provides a high degree of geometric decoupling between the X and Y motions of the motion stage resulting in large unconstrained motion range; actuator isolation that allows the use of large-stroke single-axis actuators (X Actuator and Y Actuator); and a complementary end-point sensing using commonly available sensors (e.g., Sensor 1 and Sensor 2 for the X direction).

There are many factors that make the dynamics of such a flexure mechanism challenging. Large displacements can lead to geometric relations (or coupling) between displacements in flexure mechanics. Given their dependence on the magnitude of displacement, these relations (and their impact on system dynamics) vary with the operating point of the flexure mechanism. Furthermore, large displacements require relatively low stiffness in the motion directions and therefore lead to low natural frequencies of the first few modes. Any attempt to achieve a closed-loop bandwidth that is greater than these low natural frequencies requires a proper understanding of these low-frequency dynamics, which is complicated by the above-mentioned geometric coupling. Further, while the symmetric layouts (e.g., Fig 1-2a) help provide large range, cancel undesired motions, improve space utilization, and enhance quasi-static performance, they also result in multiple closely spaced modes that are highly sensitive to manufacturing tolerances. This results in parametric uncertainty in the system dynamics.

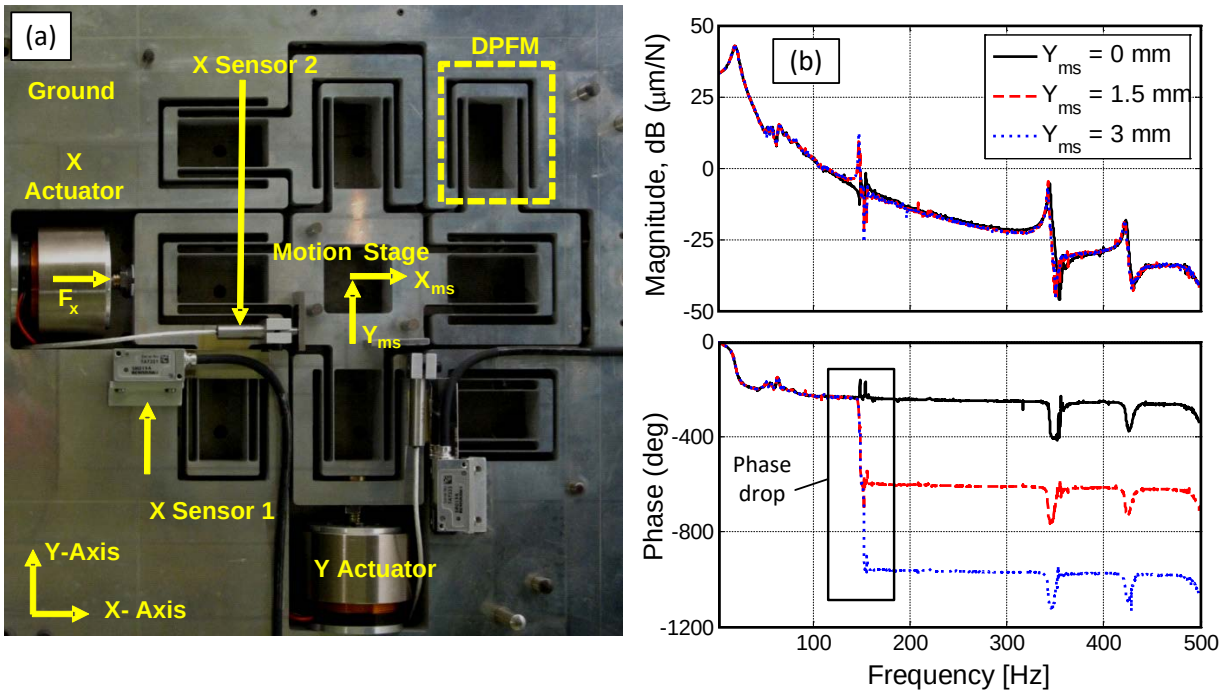

Fig 1-2 The XY flexure mechanism and its corresponding system identification: (a) the layout of the XY flexure mechanism; (b) experimental frequency response under different operation point.

Fig 1-2b shows an experimental measurement of the non-collocated X direction frequency response from force input F_x to the displacement output X_{ms}, for different values of Y

displacement (Y_{ms} = 0, 1.5, 3 mm) [11, 13]. One may notice that there are multiple closely spaced modes around 150Hz. These correspond to the natural frequency of the secondary mass in each DPFM (discussed further in Chapter 2). It is also noteworthy that the X direction frequency response changes with the Y direction operating point. At the operating points y_{ms} = 1.5 and 3 mm, the frequency response shows an additional phase drop of 360° and 720°, respectively, near 150 Hz compared to the nominal operating point y_{ms} = 0. The magnitude and phase below and above 150 Hz remains the same for all the operating points. Such observation cannot be explained by minimum phase zero pairs. Thus, such phase drop is due to CNMP zero pairs on the right half plane. This dynamic response was unexpected and the existence and number of CNMP zeros seemed arbitrary.

From a closed-loop performance stand-point, it is well known that NMP zeros severely limit bandwidth, stability robustness, and positioning speed [14, 15]. To explain this limit, here we provide a visual explanation of a key theorem that is related to the sensitivity peak of a system that has NMP zeros [16].

Suppose that $z=x+jy$ $(x>0)$ where z is a zero on the right half plane (RHP). Assume that the sensitivity function S is required to be small over a frequency range $[0, \omega_0)$:

$$\left|S\left(j\omega\right)\right| < \alpha < 1, \forall \omega \in \Omega \stackrel{\Delta}{=} \left[0, \omega_0\right) \tag{1-1}$$

Due to its role in the sensitivity function, ω_0 is also closely related to the crossover frequency and thus the closed loop bandwidth. One can assume an approximately linear relation between ω_0 and the bandwidth.

Next, we state the weighted length of the frequency interval Ω by

$$\Theta\left(z,\Omega\right) = \int_0^{\omega_0} W\left(z,\omega\right) d\omega = \int_0^{\omega_0} \left(\frac{x}{x^2 + \left(y-\omega\right)^2} + \frac{x}{x^2 + \left(y+\omega\right)^2} \right) d\omega = -\angle\left(\frac{z - j\omega_0}{z + j\omega_0} \right) \tag{1-2}$$

$\Theta(z,\Omega)$ is illustrated in Fig 1-3, which is the opening angle from the right half plane zero z, to +/- $j\omega_0$ on the imaginary axis. Based on this geometric construction, there are two main observations. First, $\Theta(z,\Omega)$ is always in the range of $[0, \pi]$. Particularly, $\Theta(z,\Omega) \rightarrow 0$ as $\omega_0 \rightarrow 0$, and $\Theta(z,\Omega) \rightarrow \pi$ as $\omega_0 \rightarrow \infty$. Second, if one draws a circle across the three points: z, $(0, j\omega_0)$, and $(0, -$

$j\omega_0$), then any RHP zero z lying on this circle has the same value of Θ. For the example in Fig 1-3a and Fig 1-3b, z_2 is another RHP zero different from z, but $\Theta(z_2,\Omega)= \Theta(z,\Omega)$. Note that if this circle is centered at the origin (Fig 1-3a), then $\Theta(z,\Omega)=\pi/2$ (i.e., 90°); in contrast, if the circle is centered on the right side of the real axis (Fig 1-3b), then $\Theta(z,\Omega)<\pi/2$.

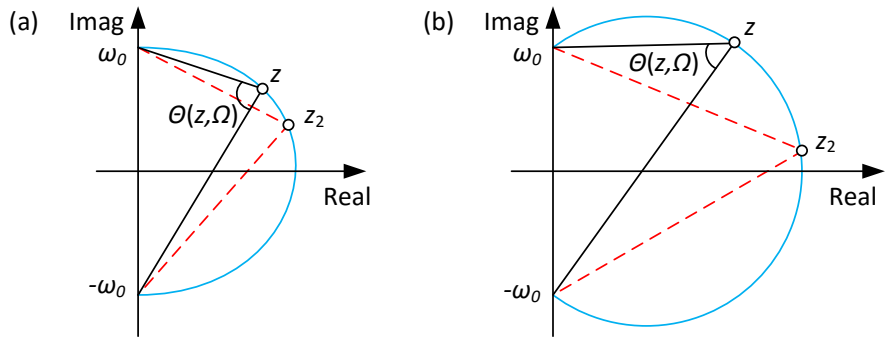

Fig 1-3 Geometric illustration of the opening angle: (a) the circle is centered at the origin; (b) the circle is centered on the right side of the real axis.

Next, the theorem [16] states that if both the open loop system (i.e., the plant with the controller) and the closed loop system are stable, and the sensitivity function is constrained by Eq.(1-1), then

$$\max_{\omega\geq\omega_0}\left|S\left(j\omega\right)\right|>\left(\frac{1}{\alpha}\right)^{\frac{\Theta(z,\Omega)}{\pi-\Theta(z,\Omega)}}\triangleq\left(\frac{1}{\alpha}\right)^{\Lambda} \tag{1-3}$$

To explain, consider first the power order term Λ. Λ is a monotonously increasing function of Θ, with a range as [0, ∞). Two main observations are provided. First, an important case is when $\Theta=\pi/2$, one has $\Lambda=1$. This corresponds to the situation in which the RHP zero of interest is on the circle that is centered at the origin and has a radius of ω_0 (Fig 1-3a). Second, if the position of the RHP zero is fixed, then increasing ω_0 results in a larger Θ, and thus a larger Λ.

With the explanation of Λ, we next consider Eq. (1-3), which is known as the guaranteed peak or maximum value of the sensitivity function. On one hand, one generally prefers a small α (i.e., small in sensitivity function at low frequency) for better disturbance rejection, smaller tracking error, etc. On the other hand, one also prefers a large ω_0 for higher bandwidth. However,

from Eq. (1-3), one can see that either decreasing α or increasing ω_0 significantly increases the maximum value of the sensitivity function. As a result, the modulus margin is significantly decreased, where modulus margin is defined as the minimum distance from the point (-1, 0) to the open loop transfer function of a system in a Nyquist plot. This indicates that the robustness of the closed loop system drops significantly. In other words, if a system has an RHP zero, then in order to maintain a reasonable robustness, one has to choose a less aggressive pair of ω_0 and α.

The above mentioned conclusion is visualized in Fig 1-4. Assume that a system has an RHP zero at z and choose α=0.5. Moreover, Θ has to be small in order to keep the sensitivity peak small. For example, if we choose $\Theta=\pi/6$ (i.e., 30°), then by drawing an opening angle of $\pi/6$ from z to the imaginary axis, ω_0 is determined. Note that if one draws a circle passing through z, (0, ω_0), and (0, -ω_0), any RHP zero on this circle will result in the same opening angle of $\pi/6$. Furthermore, if the system's RHP zero is located outside the circle (shaded area) instead of z, then using the current ω_0 will result in an opening angle smaller than $\pi/6$, which reduces the lower bound of the sensitivity peak. However, if the system's RHP zero is located inside the circle (e.g., z_2), then in order to maintain the same value of the opening angle, a smaller circle (dashed line) is required, resulting in a smaller ω_0, and therefore a lower bandwidth.

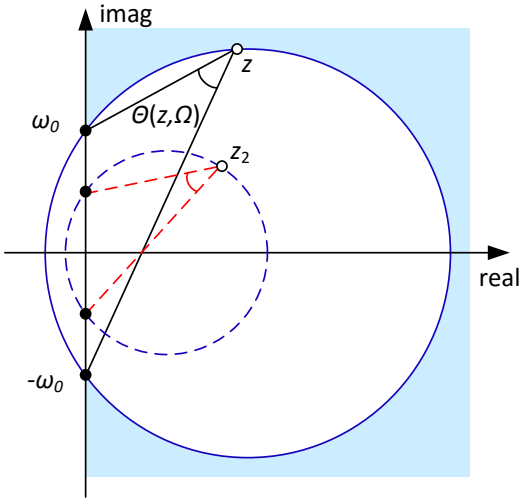

Fig 1-4 Geometric illustration of NMP zeros' limitation

From Eq. (1-3) it is preferable to have Θ smaller than $\pi/2$. Thus, the corresponding circle is centered on the real axis of the RHP. For the circle in Fig 1-4, if one examines the distance from

the origin to the circle, the minimum value is achieved on the imaginary axis, while the maximum value is achieved on the real axis. Recall that for a zero (e.g., z in Fig 1-4), its corresponding frequency is the distance from the origin to its position in the complex plane. Therefore, for RHP zeros with the same frequencies, in order to achieve the same Θ for robustness, ω_0 is the smallest when the zero is on the real axis (i.e., real NMP zero), while ω_0 is the largest when the zero is close to the imaginary axis (i.e., complex NMP zero). In other words, among the RHP zeros with the same frequencies, the real NMP zero has the strongest limitation on closed loop bandwidth.

With the above observations, several questions arise: since complex NMP zeros always exist in the form of conjugate pairs, how should a pair of complex conjugate NMP zeros be evaluated using Fig 1-4? With the experimental observation of CNMP zeros in Fig. 1-2b, when, why, and how do these CNMP zeros appear? Can they be analytically predicted? Do they have a physical meaning? Can they be avoided via physical system design? If they are mitigated, can one achieve a better dynamics performance? All of these questions represent a gap in the current knowledge of flexure mechanism dynamics. This gap, in turn, limits improvements in the performance of flexure based motion systems.

To address this knowledge gap, the goal of this Ph.D. thesis research is to develop and validate models to explain the conditions (when), the means (how), and the physical interpretations (why) of CNMP zeros that appear in flexure mechanisms. This understanding will result in guidelines for mechanical and control system design to address such CNMP behavior in flexure based motion systems. For a representative study, this research will focus on flexure mechanisms based on the double-parallelogram flexure module (DPFM) (Fig 1-2a), but its outcomes are more broadly relevant to the dynamics of flexible systems.

1.3. Problem Statement

Based on the above discussion, the problem is stated as follows. *"Conduct a scientific investigation into the existence, genesis, and consequences of CNMP zeros observed experimentally in the dynamics of large displacement flexure mechanisms."*

The scientific knowledge gained through this research is essential to overcome traditional tradeoffs between large displacement and good dynamic performance in various applications that employ flexure mechanisms for motion guidance.

1.4. Organization of the Thesis

Based on the problem statement, the work in this thesis is divided into five research tasks as below. The organization of the chapters is based on these tasks. An illustration of the research framework and task chain is presented in Fig 1-5.

Task 1 (Chapter 2): Develop reduced-order analytical models of flexure mechanisms taking into consideration the relevant geometric non-linearities due to large deformation as well as unintended parametric asymmetries (e.g., manufacturing imperfections). Predict when these two factors result in CNMP zeros in the input-output transfer function of the flexure mechanism over its range of motion.

Task 2 (Chapter 3): Conduct a systematic investigation based on modal theory to explain the existence and genesis of CNMP zeros in systems that have two closely spaced modes. Use the analytical results to generate a physical interpretation of the CNMP zeros.

Task 3 (Chapter 4): Explore the possibility of an inherent relationship between the appearance of CNMP zeros and the phenomenon of curve veering seen in systems with symmetric/periodic structures. This relationship has not been explored before and, if established, would enable a deeper understanding of both CNMP zeros and curve veering.

Task 4 (Chapter 5): Design and conduct experiments to validate the appearance of CNMP zeros as predicted for any given operating point within the range of motion and parametric asymmetry. This experimental set-up will comprise a flexure mechanism integrated with sensors, actuators, and drivers, as well as control hardware and software.

Task 5 (Chapter 6): Based on the above physical system understanding, design mechatronic systems (i.e., mechanical system and controller) to demonstrate superior dynamic performance over large motion range.

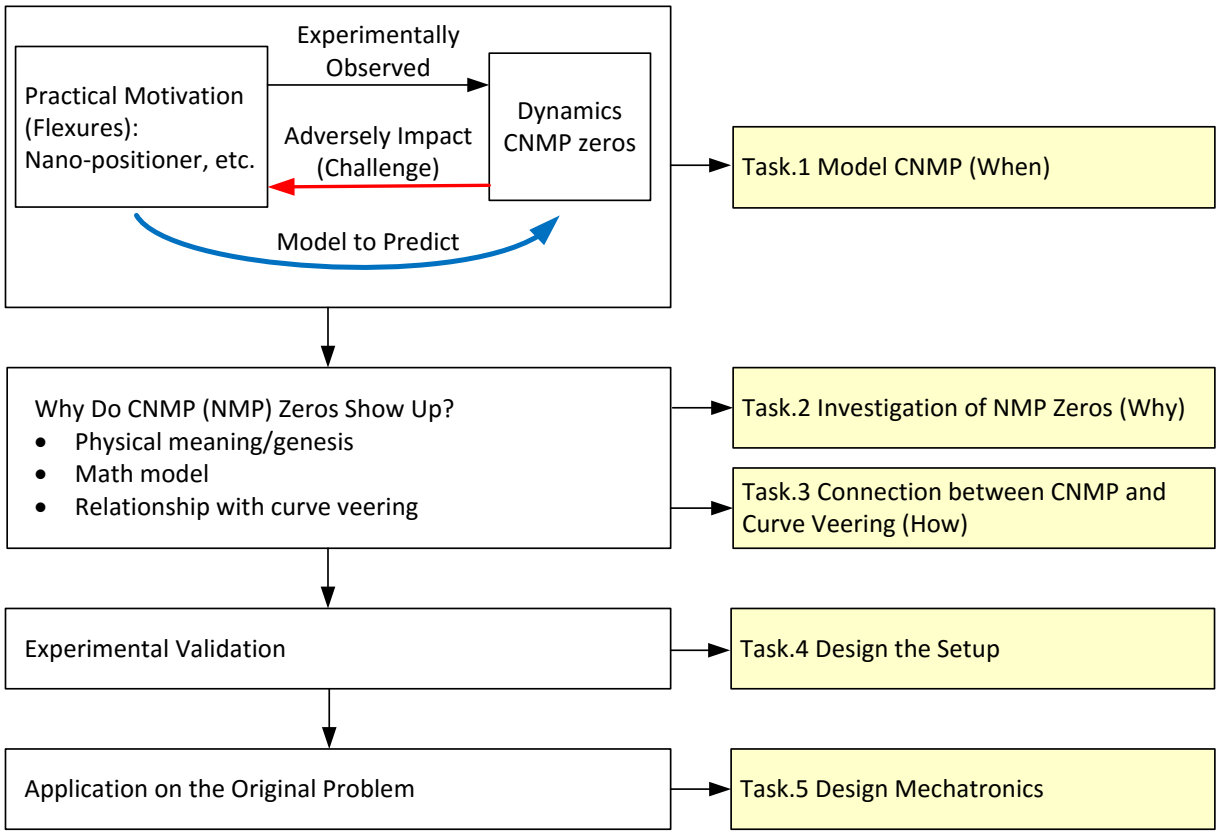

Fig 1-5 Illustration of research framework and related tasks

1.5. Research Contributions

The contributions of this thesis are summarized as below:

(1) A lumped-parameter modeling approach that analytically captures the dynamics of flexure mechanisms comprising the parallelogram or double parallelogram modules is established. This model captures the key relevant geometric non-linearity in large displacement flexure mechanics. Linearization about any arbitrary operating point enables frequency domain analysis. With this model we are able to predict previously unexplained CNMP zeros seen experimentally. The model establishes the existence of CNMP zeros under certain combination of operating point and parametric asymmetry in the non-collocated transfer function of a simple representative XY flexure mechanism.

(2) A theoretical model is presented to explain the conditions (when), the means (how), and the physical interpretations (why) of the occurrence of the CNMP zeros in flexure mechanisms. For NMP zeros, the well-known limitations on control are visualized in a Nyquist plot.

(3) The connection between CNMP zeros and curve veering is explored. Using normalization and the variation of modal residues, it is shown under certain conditions that CNMP zeros are guaranteed in the presence of curve veering. This finding enables a further understanding on both CNMP zeros and curve veering.

(4) The above modeling work is validated experimentally. The experiment is carefully designed to isolate the features of interest (i.e., the closely spaced modes) from other factors (e.g., parasitic dynamics). The existence of CNMP zeros are validated in lightly damped, passive stable flexure mechanisms.

(5) A mechatronic design approach (i.e., a physical and control system design performed in conjunction) is presented that leverages the above system-level knowledge to improve the dynamic performance (bandwidth, disturbance rejection, point to point, robustness, etc.) of large range flexure-based motion stages over the state-of-the-art stages.

Chapter 2. Modeling CNMP Zeros of DPFM based Flexure Mechanisms

This chapter presents a model to explain complex non-minimum phase (CNMP) zeros seen in the non-collocated frequency response of a large displacement XY flexure mechanism, which employs multiple double parallelogram flexure modules (DPFM) as building-blocks. Geometric non-linearities associated with large displacement along with the kinematic under-constraint in the DPFM, lead to a coupling between the X and Y direction displacements. Via a lumped-parameter model that captures the most relevant geometric non-linearity, it is shown that specific combinations of the operating point (i.e. flexure displacement) and mass asymmetry (due to manufacturing tolerances) give rise to CNMP zeros. This model demonstrates the characteristics that an intentionally positive/negative asymmetric configuration can guarantee/avoid the appearance of CNMP zeros, respectively. Furthermore, a study of how the eigenvalues and eigenvectors of the flexure mechanism vary with the operating point and mass asymmetry indicates the presence of curve veering when the system migrates from minimum phase to CNMP. Based on this, the hypothesis of an inherent correlation between CNMP zeros and curve veering is proposed (the hypothesis is discussed in Chapter 4).

2.1. Literature Review

In recent years, there has been a growing body of research literature on the dynamics of flexure mechanisms. Lan et al. [17] presented a distributed-parameter dynamic modeling approach of elastic flexure mechanisms. The resulting equations of motion in the time domain were solved using numerical methods, which are not very suitable for frequency domain analysis of CNMP zeros. Akano and Fakinlede [18] used finite element based non-linear analysis to predict the effect of design parameters on the dynamic performance of flexure mechanisms. While accurate, these methods are computationally intensive and provide limited physical

insights in the frequency domain. Alternatively, lumped-parameters closed-form modeling approaches have also been investigated. Shilpiekandula and Youcef-Toumi [19] derived a lumped-parameter dynamic model of a diaphragm flexure using Timoshenko beam theory, but did not include geometric non-linearities. Awtar and Parmar [11] captured the non-linear variation in the stiffness of flexure building blocks to create a lumped-parameter dynamic model of a XY flexure mechanism (Fig 1-2a), but did not capture the non-linear coupling between X and Y directions in a DPFM and therefore were unable to predict the NMP behavior seen experimentally (Fig 1-2b). The pseudo-rigid-body approach has also been used for modeling the non-linear dynamics of flexure mechanisms [20-22]. While this approach leads to simple lumped-parameter closed-form models, the model parameters are computed via numerical optimization and depend on the boundary conditions of each beam, thereby increasing the modeling complexity in flexure mechanisms that have a large number of beams.

Dynamic modeling of rigid link mechanisms with inherent flexibilities, e.g. robotic manipulators, has also been an active area of research. An overview and classification of various modeling approaches is found in the review paper by Dwivedy and Eberhard [23]. Research in this area includes the study of manipulators with one or more flexible links as well as one or more flexible joints. Various methods including finite elements, assumed modes, lumped parameter, and inverse dynamics have been adopted to study the relevant dynamics. This body of work assumes small deformation of the links, compared to rigid body motion, which is justified since the links are designed to be stiff. However, this assumption fails for flexure mechanisms that provide large deformation in their motion directions.

The large deformation of constituent elements or beams in a flexure mechanism results in geometric non-linearities arising from arc-length conservation, cross-sectional warping, trapeze effect, and Wagner's effect in beam mechanics [24-30]. The impact of these non-linearities on the dynamics of flexible beams and structures has been studied extensively, as reported in the review papers by Modi [31] and Pandalai [32]. Furthermore, the dynamics and control of flexible beams with an end-mass [33] as well as rotating beams [34] have also been investigated. In further generalization, DaSilva formulated the non-linear differential equations of motion for Euler-Bernoulli beams experiencing flexure along two principal directions, along with torsion and extension [35]. Jonker has formulated a highly generalized model for spatial beams taking

into account relevant non-linearities, using finite element based multi-body dynamics computations [26, 36]. Nayfeh modeled the non-linear transverse vibration of beams with properties that vary along the length [37]. Zavodney and Nayfeh studied the non-linear response of a slender beam with a tip mass to a principal parametric excitation [38]. Moeenfard studied the in-plane flexural and axial vibration of a flexure beam with a tip mass while accounting for the non-linearity associated with arc length conservation [39]. While the resulting non-linear equations of dynamics are solved in time-domain via perturbation, homotopy, or computational methods, this prior work [32] does not pursue the frequency domain investigation relevant to the present work.

Separately, there exists a significant body of work in the frequency domain dynamics of lumped or distributed parameter flexible systems [40-42]. It has been shown that lightly damped flexible systems with collocated sensor and actuator have alternating poles and zeros along the imaginary axis and are easy to stabilize in closed-loop [43, 44]. Non-collocated systems do not share these attributes and, under certain conditions, exhibit real non-minimum phase (NMP) zeros in the right half plane [45-47]. Spector and Flashner [45] studied the sensitivities of beam cross-section, material properties, and sensor placement on the locations of poles and zeros in flexible systems. They showed that as the sensor placement is moved away from the actuator, the conjugate zeros, originally located along the imaginary axis, migrate towards infinity and then reappear along the real axis. Miu [46] provided a physical explanation for these real NMP zeros stating that they are related to the non-propagation of energy within the structural subsystem confined by the actuator and sensor. Unlike real NMP zeros, CNMP zeros are relatively rare and have been reported in the context of a non-collocated acoustical transfer function of a room [48], as well as in a non-collocated transfer function of a lumped parameter spring-mass system [49, 50]. Awtar identified CNMP zeros arising due to an electromagnetic coupling between a DC motor and tachometer used in a servo-system [51]. These studies on CNMP zeros simply report a mathematical or experimental observation, without providing further insight into when or why the zeros appear.

2.2. Modeling Geometric Non-linearity in DPFM

A DPFM, shown in Fig 2-1, comprises a reference stage, a secondary stage, and a primary stage connected via two parallelogram flexures in series. This arrangement uses geometric reversal to cancel the X direction kinematic error motion (i.e. parabolic trajectory) of one parallelogram with that of the second parallelogram. In quasi-static operation, these two error motions exactly cancel out, resulting in a straight-line motion along the Y direction between the primary and reference stages while providing high stiffness in the X and Θ axes. Furthermore, the DPFM provides large displacement range and low stiffness in the Y direction for a given footprint. For a building block, these are desirable attributes because they help minimize various error motions and enable large motion range in the resultant flexure mechanisms [11].

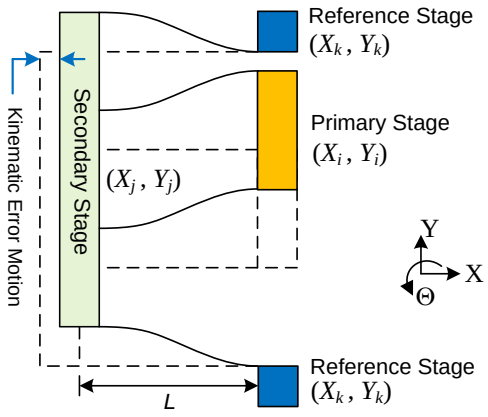

Fig 2-1 Double parallelogram flexure module (DPFM)

One of the limitations of the DPFM is that it presents a kinematic under-constraint associated with the Y displacement of its secondary stage. In quasi-static operation, this under-constraint adversely impacts the X direction stiffness, as reported previously [30]. In dynamic operation, this leads to an additional degree-of-freedom (DoF) in the DPFM even though it is intended to be a single-DoF building block. Referring to the X and Y displacements of the primary and secondary stages with respect to the reference stage in Fig 2-1, the following quasi-static relations hold [13, 30]:

$$X_i - X_j = -0.6\left(Y_i - Y_j\right)^2 / L \tag{2-1}$$

$$X_j - X_k = 0.6\left(Y_j - Y_k\right)^2 / L \tag{2-2}$$

Here, L stands for the beam length (Fig 2-1). These kinematic relations, which define the X direction error motion of the parallelogram flexures, are the result of the geometric non-linearity associated with beam arc-length conservation. Relative to the reference stage, there are four displacement coordinates in the DPFM (i.e. X_i, Y_i, X_j, Y_j); of these, only two are independent because of the above kinematic relations. This illustrates the 2 DoFs of the DPFM and its under-constraint.

The above quasi-static relations are based on certain core assumptions that may be extended to low-frequency dynamics spanning the resonance of the secondary stage of a DPFM with respect to the reference and primary stages. In the flexure mechanism of Fig 1-2a, this is the frequency range in which complex non-minimum phase zeros appear. Since the secondary stage mass is small compared to the other stages and the X direction stiffness of the beams is relatively high, modes associated with beam stretching appear at much higher frequencies. This observation justifies the first assumption that the individual flexure beams may be treated as inextensible. Beam inextensibility also eliminates Θ displacement coordinates in the DFPM. Furthermore, in the frequency range of interest, the beams are also assumed to be massless, given the relatively larger masses of the various stages. This implies that the beams deform in the quasi-static S-shape with amplitude dictated by the relative Y displacement of the relevant parallelogram flexure [29, 30]. At higher frequencies, there will be other beam shapes dictated by the resonance of individual beams, in which case the above relations will no longer hold.

Because of large out-of-plane stiffness, the out-of-plane displacements (and associated modes) are neglected in the dynamic modeling. Also, damping is neglected since the flexure mechanism is monolithic with no sliding joints. Experiments confirm that the dissipation from material hysteresis and viscous effects is very low (damping ratios about 0.001). Beyond arc-length conservation, there are several other non-linearities in beam mechanics [27, 28], but these are systematically estimated and neglected using an order of magnitude analysis [13].

2.2.1. Dynamic Modeling of XY Mechanisms[1]

The kinematic non-linearity of Eqs. (2-1) and (2-2) may be incorporated in deriving the equations of motion for a flexure mechanism involving the parallelogram or double parallelogram flexure modules. To investigate how the frequency domain dynamic response varies with the operating point, these non-linear relations may be linearized about an arbitrary operating point. But this potentially holds the risk of premature linearization. To test this possibility, we retain the non-linear kinematic relations throughout the derivation of the dynamic equations of motion and linearize the latter in the end. The results prove to be the same as when the kinematic relations themselves are linearized at the onset.

Thus, defining $Y_{ij} = Y_i - Y_j$, using the subscript "o" to denote nominal values at an operating point, and lower case letters to denote deviation from these nominal values, Eq. (2-1) becomes:

$$X_{io} + x_i - \left(X_{jo} + x_j \right) = -0.6 \left(Y_{ijo} + y_{ij} \right)^2 / L \tag{2-3}$$

Since the nominal values are still related by Eq. (2-1), Eq.(2-3) may be linearized for small deviations about the nominal values

$$x_i - x_j = -1.2 \, (Y_{ijo} / L) \, y_{ij} \overset{\Delta}{=} -\alpha \cdot y_{ij} \tag{2-4}$$

Here α is a *coupling coefficient* that depends on the operating point Y_{ijo} and captures the coupling between the X and Y axis displacement coordinates.

Next, for the purpose of investigating CNMP zeros in XY flexure mechanisms that have multiple DPFM building blocks, we select a simple representative flexure mechanism shown in Fig 2-2 to initially limit modeling complexity and enable physical insights into the observed dynamic phenomena. Yet, this design captures all the essential attributes of more complex flexure mechanisms (e.g. Fig 1-2a). The layout comprises the smallest number of symmetrically placed DPFMs (i.e. two) needed to produce multiple (i.e. two) closely spaced modes associated with the kinematic under-constraint of the secondary stages (② and ③). This mechanism allows stage ① to displace in the X and Y directions. The former is due to the X direction

[1] This work was done in collaboration with Gaurav Parmar (M.E. Ph.D. 2014) at the University of Michigan

bearings (indicated by the rollers) at stages ④ and ⑤ and the latter is due to large bending deformations of the beams in the two DPFM. Large deformation leads to the geometric non-linearity and associated coupling between X and Y displacements, mentioned above. For this mechanism, the operating point is given by a static displacement of stage ① in the Y direction with respect to ground (Y_{1o}) caused by a constant force F_{y1o}. Thus, the non-collocated transfer from an X direction force P on stage ④ to the X displacement of stage ① (X_1) can be investigated for different values of Y_{1o}. In fact, this simple mechanism is representative of a portion (indicated by the red rectangular outline) of the more complex XY mechanism of Fig 1-2a.

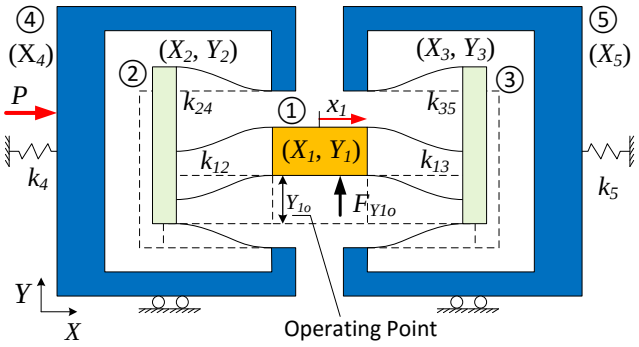

Fig 2-2 Simple representative flexure mechanism

The five stages have eight displacement coordinates, as shown in Fig 2-2; lower case versions of these coordinates represent the respective deviations with respect to an operating point. Furthermore, these coordinates are related by four kinematic relations (Eq.(2-4)), one for each parallelogram. Thus, this mechanism has four DoF and is therefore referred to as the simple representative 4-DoF (SR4DoF) mechanism in this thesis. The displacement coordinates x_1, y_1, y_2, and y_3 are chosen for this analysis, to be able to study the X displacement of stage ① and the Y displacements of stages ② and ③ with respect to stage ①.

Assuming a lumped parameter Y direction bending stiffness for each of the parallelograms (k_{12}, k_{24}, k_{13}, k_{35}), the equations of motion for the SR4DoF may be derived to be the following:

$$M\,\ddot{\mathbf{z}} + K\,\mathbf{z} = \mathbf{Q} \tag{2-5}$$

$$\mathbf{z} = [x_1 \quad y_1 \quad y_2 \quad y_3]^T \quad \text{and} \quad \mathbf{Q} = [1 \quad \alpha \quad -2\alpha \quad 0]^T \cdot P$$

$$M = \begin{bmatrix} m_1+m_2+m_3+m_4+m_5 & \alpha(m_2+m_4-m_3-m_5) & -\alpha(m_2+2m_4) & \alpha(m_3+2m_5) \\ \alpha(m_2+m_4-m_3-m_5) & m_1+\alpha^2(m_2+m_3+m_4+m_5) & -\alpha^2(m_2+2m_4) & -\alpha^2(m_3+2m_5) \\ -\alpha(m_2+2m_4) & -\alpha^2(m_2+2m_4) & m_2+\alpha^2 m_2+4\alpha^2 m_4 & 0 \\ \alpha(m_3+2m_5) & -\alpha^2(m_3+2m_5) & 0 & m_3+\alpha^2 m_3+4\alpha^2 m_5 \end{bmatrix}$$

$$K = \begin{bmatrix} k_4+k_5 & \alpha(k_4-k_5) & -2\alpha k_4 & 2\alpha k_5 \\ \alpha(k_4-k_5) & k_{12}+k_{13}+\alpha^2(k_4+k_5) & -2\alpha^2 k_4-k_{12} & -2\alpha^2 k_5-k_{13} \\ -2\alpha k_4 & -2\alpha^2 k_4-k_{12} & 4\alpha^2 k_4+k_{24}+k_{12} & 0 \\ 2\alpha k_5 & -2\alpha^2 k_5-k_{13} & 0 & 4\alpha^2 k_5+k_{13}+k_{35} \end{bmatrix}$$

It should be noted that the lower-case displacement coordinates in the above equations of motion represent the small deviations about the respective nominal operating point values. Based on these equations, one can derive the transfer function *G(s)* from the input force *P* to output displacement x_1 for different values of α, which depends on Y_{1o}:

$$\alpha = 0.6\ Y_{1o} / L \tag{2-6}$$

In the above stiffness matrix, the Y direction bending stiffness for each parallelogram (k_{12}, k_{24}, k_{13}, k_{35}) is nominally $24EI/L^3$, where E is the bending modulus, L is the beam length, and I is the second moment of area about the Z axis [30]. To obtain numerical results, we use the same dimensions as those for the XY flexure mechanism of Fig 1-2a (see Table 2-1). Furthermore, although not included in the above derivation, small nominal values for damping are assumed to avoid singularities in the numerical simulation. Although the SR4DoF is intended to be symmetric, there is the possibility for parametric asymmetry between (m_4 and m_5), (m_2 and m_3), (k_4 and k_5), or (k_{12}, k_{24}, k_{13}, and k_{35}) resulting from finite manufacturing tolerances. Initially, we assume the parameters be perfectly symmetric; but the above lumped parameter model allows us to study the impact of asymmetries in the next section.

The last three of the four predicted modes are shown in Fig 2-3 (relative to the displaced configuration / operating point of Fig 2-2) while all modes are quantified in Table 2-2. There are three key observations: (1) The 1st mode (not shown in Fig 2-3) is the 'rigid body' mode in which all the stages vibrate together in the X direction due to springs k_4 and k_5; (2) The 2nd mode is associated with the in-phase Y direction vibrations of stages ①, ②, and ③; (3) The 3rd and 4th modes are primarily associated with the Y direction vibrations of the two secondary stages

(② and ③) in the opposite and same directions, respectively, with natural frequencies close to each other.

Table 2-1. Dimensions and physical parameters.

Name	Symbol	Value	Unit
Beam length	L	47.5	mm
Beam height	H	25	mm
Beam thickness	T	0.625	mm
Young's modulus (aluminum)	E	6.9×10^{10}	N/m^2
Second moment of area	I	5.09×10^{-4}	mm^4
Poisson's ratio	v	0.33	
Motion stage mass	m_1	0.284	kg
Secondary stage mass	m_2, m_3	0.018	kg
Intermediate stage mass	m_4, m_5	0.177	kg
Spring stiffness	k_4, k_5	4.4×10^3	N/m

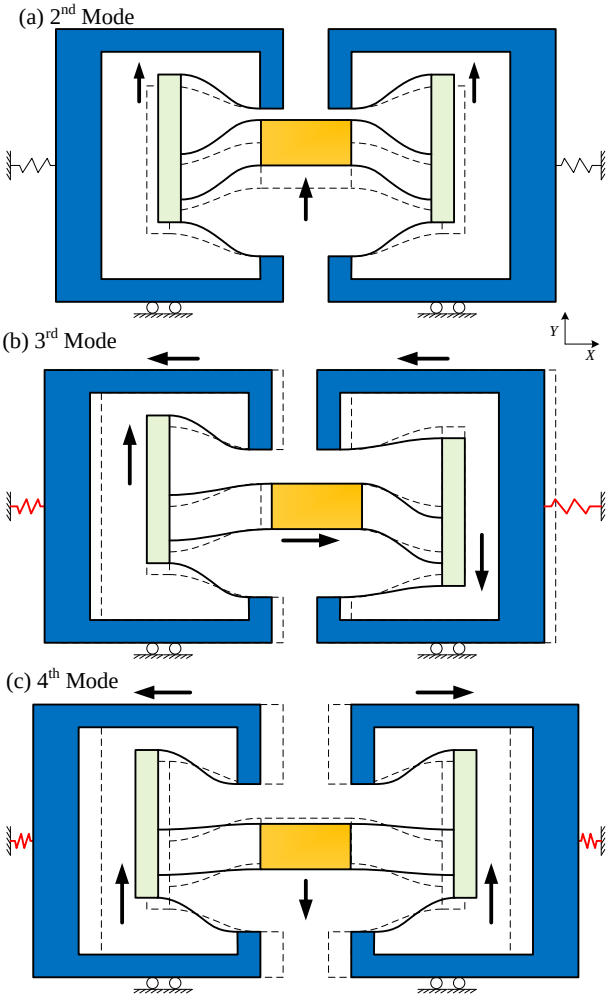

Fig 2-3 Mode shapes of the SR4DoF. Black arrows indicate the relative motion direction of each stage.

Table 2-2. Eigenvectors of the SR4DoF mechanism at operating point Y_{1o}=5% of *L*.

Δm_{23}=0%	f_1=18.2 Hz	f_2=27.6 Hz	f_3=153.2 Hz	f_4=154.3 Hz	Δm_{23}=5%	f_1=18.2 Hz	f_2=27.6 Hz	f_3=150.1 Hz	f_4=153.9 Hz
x_1	0.9999	0	-0.0231	0	x_1	0.9999	-0.0000	-0.0189	-0.0139
y_1	0	0.8075	0	0.0481	y_1	0.0000	0.8073	0.0294	-0.0387
y_2	0.0068	0.4171	-0.7069	-0.7063	y_2	-0.0068	0.4176	-0.9845	0.1384
y_3	-0.0068	0.4171	0.7069	-0.7063	y_3	0.0068	0.4169	0.1718	0.9895

The 3rd and 4th modes arise due to the under-constrained secondary stages in the DPFM. When $Y_{1o} = 0$ (i.e. $\alpha = 0$), the vibration of the secondary stages does not cause any X direction motion of stage ①. Therefore, these two modes are unobservable in the *G(s)* transfer function. However, when $Y_{1o} \neq 0$, the X and Y displacements of the DPFMs get coupled, which affects the 3rd and 4th modes differently. For the 3rd mode, the Y vibration of the two secondary stages is coupled to the X vibration of stage ①. However, in the 4th mode, the two secondary stages have the same vibration magnitude and phase in the Y direction, which results in a cancellation of the coupling at stage ① in the X direction. Instead, the coupling results in X direction vibrations at stages ④ and ⑤ (Fig 2-3c). Thus, when parameters are symmetric but $Y_{1o} \neq 0$, the 3rd mode shows up in the *G(s)* transfer function while the 4th mode remains unobservable.

2.2.2. Parametric Asymmetry and CNMP Zeros

While the SR4DoF was assumed to be perfectly symmetric so far, some parametric asymmetry is inevitable in a practical situation due to manufacturing tolerances. Using the above lumped parameter model, we varied parametric asymmetry over a ± 5% range for (m_4 and m_5) and (k_4 and k_5). This did not have much of an effect on the flexure mechanism dynamics in terms of mode shapes and transfer functions. But an asymmetry in (k_{12}, k_{24}, k_{13}, and k_{35}) or (m_2 and m_3) impacts the vibrations of the secondary stages (i.e. 3rd and 4th modes) and therefore the overall flexure mechanism. Of the two sets of parameters, the mass parameter is more sensitive. When the DPFM is used as a building block, its secondary stage size and mass are minimized to reduce

footprint and raise the resonance frequency at which it vibrates. For example, in the XY flexure mechanisms of Fig 1-2a and Fig 2-2, the nominal mass of the secondary stage is 18g. Therefore, even a small additional mass such as 0.9g results in a relatively large variation (5%). Therefore, in this section, we investigate how an asymmetry in masses, Δm_{23} (= $m_2/m_3 - 1$), affects the dynamics of SR4DoF.

As seen via the respective eigenvectors in Table 2-2, the impact of non-zero Δm_{23} on the 1st and 2nd modes is minimal. This mass asymmetry primarily impacts the vibration of the two secondary stages, which directly influence the 3rd and 4th modes. As noted earlier, the 4th mode of the SR4DoF flexure mechanism is hidden in the transfer function *G(s)* for any Y_{1o} (zero or non-zero) when $\Delta m_{23} = 0$. However, for a small parametric asymmetry, e.g. $\Delta m_{23} = 5\%$, the two secondary stages have different vibration magnitudes as seen in the eigenvector in Table 2-2. Thus, the X direction coupling no longer cancels out at stage ① and the 4th mode appears in *G(s)*. Similarly, the impact of $\Delta m_{23} = 5\%$ on the 3rd mode is significant.

G(s) is plotted in Fig 2-4 as Y_{1o} varies from 0% to 5%, for Δm_{23} = 0% and 5%. Key observations are: (1) As expected, the 3rd mode is hidden when $\Delta m_{23} = 0$ but appears when Δm_{23} = 5%; (2) The natural frequencies of the 3rd and 4th modes drop as Y_{1o} increases. As the X vibration of stage ① gets increasingly coupled with the Y vibration of the secondary stages, the modal mass increases more than the modal stiffness, resulting in reduced natural frequencies; (3) A 360° phase drop is observed at around 150 Hz in the asymmetric case ($\Delta m_{23} = 5\%$) when $Y_{1o} =$ 5% but not in a symmetric case (Δm_{23} = 0%). In the latter case, the phase drop due to the complex pole pair (3rd mode) is off-set by a phase rise due to the complex zero pair ("valley"), resulting in no net phase drop. For the asymmetric case, there are two stable complex pole pairs (3rd and 4th modes), each contributing 180° phase drop. But since the overall phase drop is 360°, this implies that there is no phase rise or drop at the "valley" even though there are two pairs of complex zeros at in this frequency region (about 153 Hz). This indicates the presence of a quartet of complex zeros, with one pair in the left half plane and the second pair in the right half plane, thereby contributing no net phase change. This is an important observation because it suggests that CNMP zeros can arise at certain combinations of operating points and parametric asymmetry.

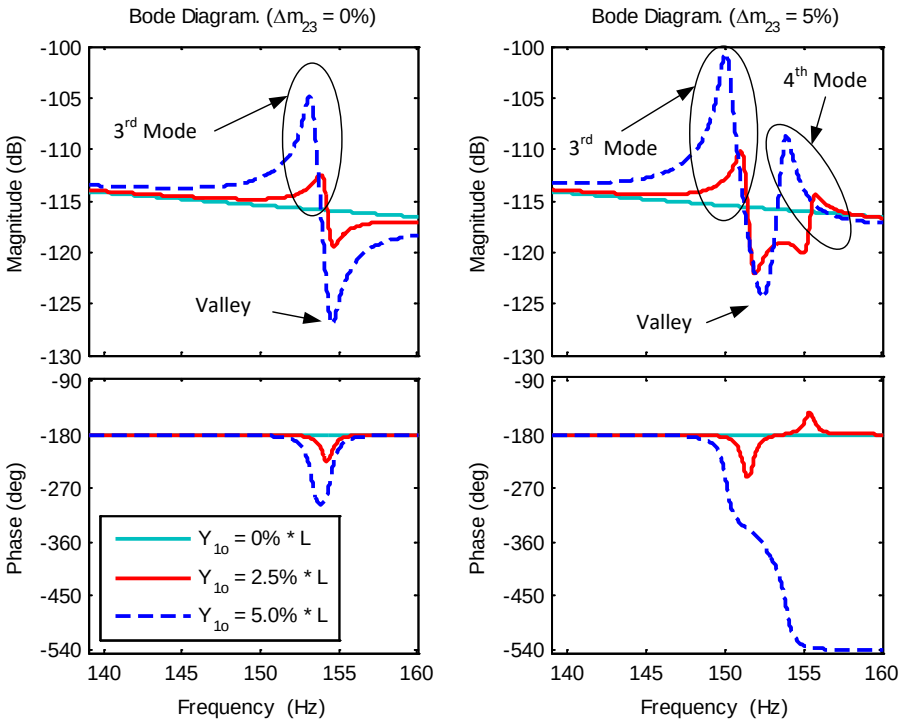

Fig 2-4 *G(s)* transfer function for different operating points (Y_{1o}) and mass asymmetry (Δm_{23}).

2.2.3. Existence of CNMP Zeros

Next, we proceed to analytically determine the conditions under which CNMP zeros arise. Based on modal decomposition, the transfer function *G(s)* may be written as follows, where β_i is the modal residue and ω_i is the corresponding natural frequency:

$$G(s) = \sum_{i=1}^{4} \frac{\beta_i}{s^2 + \omega_i^2} \overset{u=s^2}{=} \frac{b_3 u^3 + b_2 u^2 + b_1 u + b_0}{a_4 u^4 + a_3 u^3 + a_2 u^2 + a_1 u + a_0} = \frac{N(u)}{D(u)} \tag{2-7}$$

The decomposed form can also be expressed via a numerator *N(u)* and a denominator *D(u)*, each a polynomial. Note that there are no odd power *s* terms because damping is ignored. If *N(u)*, which is a cubic polynomial, has two complex conjugate roots, then *G(s)* will have a quartet of complex zeros. Two of these zeros will be in the right half plane (i.e. CNMP zeros). For this to happen, the following inequality in the coefficients of *N(u)* has to hold [52]:

$$\Delta = 18 b_3 b_2 b_1 b_0 - 4 b_2^3 b_0 + b_2^2 b_1^2 - 4 b_3 b_1^3 - 27 b_3^2 b_0^2 < 0 \tag{2-8}$$

Therefore, this is the mathematical condition for the existence of CNMP zeros in the *G(s)* transfer function. Shown in Fig 2-5, Δ is plotted in a contour map against a range of operating points and parametric asymmetry values for the SR4DoF flexure mechanism. The color in the contour map represents the magnitude of Δ: red represents higher positive values, blue represents lower positive values, and the black region represents the conditions for which Δ becomes negative, indicating the presence of CNMP zeros. This particular mechanism is seen to be very sensitive to positive asymmetry, i.e., if m_2 is greater than m_3 even by a small amount, then CNMP zeros arise in specific ranges of Y_{1o}. However, if $m_2 < m_3$, then the entire operating range is free of CNMP zeros. The reason for such asymmetric behavior is due to the physical asymmetry introduced by the actuator placement in Fig 2-2).

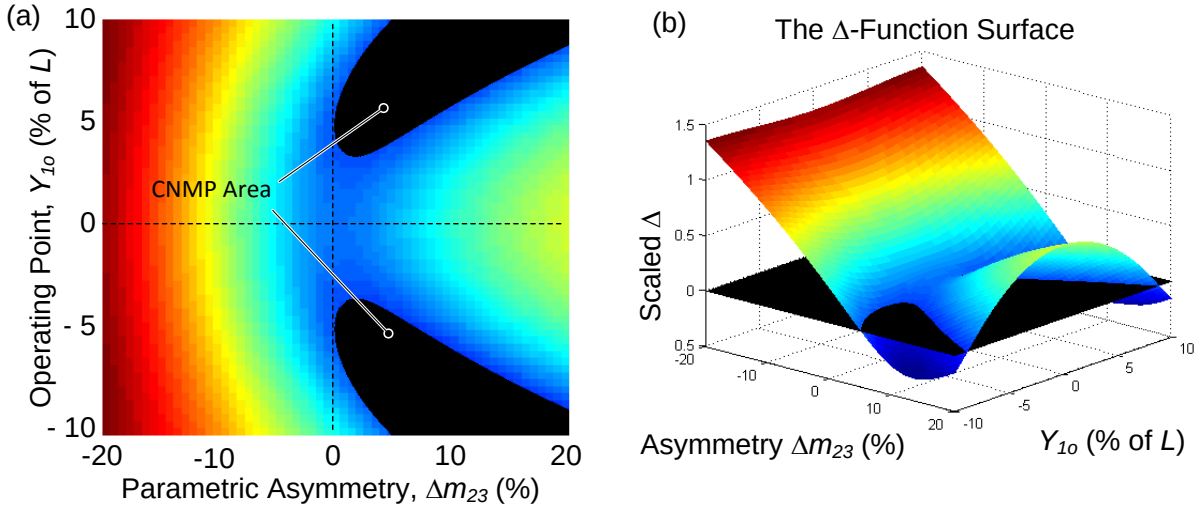

Fig 2-5 Contour map of Δ function (a) top view; (b) 3D view.

With this finding, we are able to replicate via modeling some aspects of the NMP phenomenon previously observed experimentally (Fig 1-2b). Although, for this study we intentionally chose the SR4DoF mechanism to keep modeling complexity and assumptions minimal, it is representative of the more complex designs in that it incorporates the key attributes of DPFM building blocks (with their under-constrained secondary stages), geometrically symmetric design, large displacements leading to non-linear coupling between axes, non-collocated transfer functions, and parametric asymmetry.

In the design of multi-DoF flexure mechanisms, it is a common guideline to employ symmetric and/or periodic geometries to cancel undesired motion, improve space utilization and enhance quasi-static performance [3]. However, the above dynamic model for the SR4DoF mechanism indicates that a perfectly symmetric layout is sensitive to parametric asymmetry, which is likely to occur due to manufacturing tolerances and can give rise to CNMP zeros. But if the design is intentionally made asymmetric (i.e. if Δm_{23} is sufficiently negative), then CNMP zeros can be avoided even if there are finite manufacturing tolerances.

This also shows that one can pursue the springs in the flexure mechanism (i.e. beam flexures dimensions and layout) to be symmetric to achieve the desired quasi-static performance, while choose certain masses to be asymmetric which provides the desired dynamic performance without impacting quasi-static performance. This combination of symmetry in certain attributes and asymmetry in others is rather counter-intuitive but helps meet both quasi-static and dynamic performance goals.

In Chapter 6 we will see that, with certain intentionally mass asymmetric design, we also create the possibility of achieving a higher closed-loop bandwidth, while maintaining robustness, in the SR4DoF flexure mechanism, comparing to other mass configuration. This would have been impossible without the knowledge above.

2.2.4. Modeling a Complex XY Mechanism

So far, we modeled the SR4DoF flexure mechanism to predict CNMP zeros at around 150 Hz. Next, we extend this modeling approach to the more complex (and practically relevant) XY flexure mechanism of Fig 1-2a. The transfer function X actuator force F_x to the X displacement of the motion stage X_{ms} is investigated. The Y actuator is used to provide a constant force F_{yo} to achieve various Y direction operating points Y_{mso}. There are 13 rigid stages in this case, each with an X and Y displacement coordinate. Each of the 16 parallelogram flexures provides one kinematic relation between relative X and Y coordinates. Therefore, the model has 10 independent DoFs, which results in the same number of equations of motion, natural frequencies, and mode shapes. There is a “rigid body” X vibration mode, a “rigid body” Y vibration mode,

and eight modes associated with the vibration of the secondary stages (all around 150Hz) in the eight DPFM. Next, we arbitrarily vary these secondary stage masses with respect to their nominal value, and use the model to predict the 0° / 360° / 720° phase drop seen in Fig 1-2b and the existence of CNMP zeros. The secondary stages are identified by the subscript a-h, as shown in Fig 1-2a. We present three cases in Table 2-3 with different combinations of secondary stage mass variations that result in the three different phase drops. Thus, we are able to analytically predict the seemingly unexplained phenomena observed experimentally in Fig 1-2b.

Table 2-3. Deviations of masses that cause different phase drops.

Phase Drop	Y_{ms} (% of L)	% deviation in masses							
		m_a	m_b	m_c	m_d	m_e	m_f	m_g	m_h
0	0	0	0	0	0	0	0	0	0
360°	4.88	-8.16	13.10	8.57	-13.21	-13.24	0.66	-4.80	-13.12
720°	9.54	8.47	12.08	-14.19	-12.78	6.59	-2.42	14.97	5.14

2.3. Possible Relation between CNMP Zeros and Curve Veering

As noted earlier, the SR4DoF mechanism considered in this chapter has a symmetric and repetitive geometry, which leads to two closely spaced modes (i.e., the 3rd and 4th modes shown in Table 2-3). Furthermore, these modes vary with the operating point and parametric asymmetry. The operating point determines the extent of the cross-axis coupling between X and Y displacements, thus building a connection between the 3rd and 4th modes. All these features make the phenomenon of curve veering (or mode veering) potentially relevant in this study.

Curve veering occurs when the eigenvalue loci of two closely spaced modes in a system approach each other and then diverge, as a result of parameter variation [53]. The point in the parameter space where the two modes are the closest is called the *veering point*. The special case when the two modes intersect at the veering point is called *mode crossing* [54]. In the vicinity of the veering point, eigenvectors undergo dramatic changes. As a result, the system could become so-called “critically configured” meaning that small changes in a system parameter could cause large changes in system response [55].

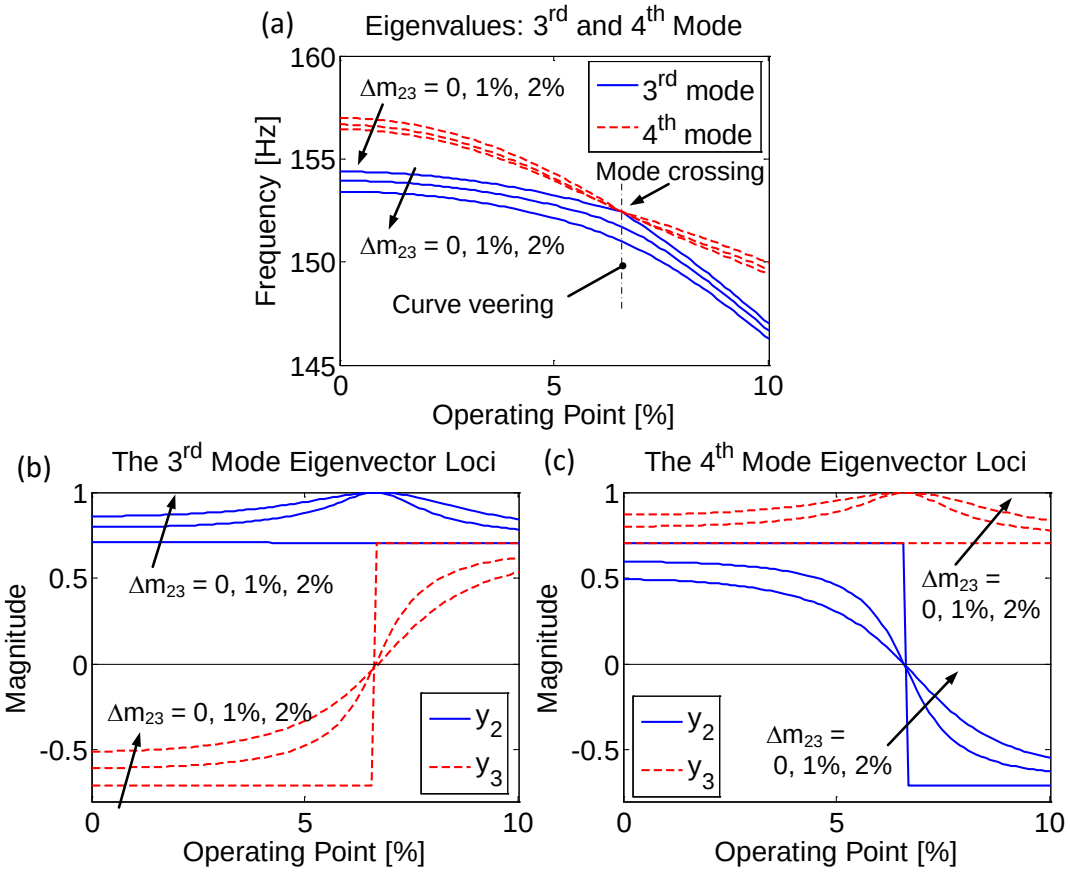

Fig 2-6 a. Eigenvalue loci of the 3rd and 4th modes; b. and c. Eigenvector elements of the 3rd and 4th modes. The sign of magnitude indicates the phase.

In the SR4DoF flexure mechanism dynamics, we observe curve veering close to the operating point and parametric asymmetry value at which the CNMP zeros arise. Fig 2-6a shows the eigenvalue loci of the 3rd and 4th modes as a function of the operating point (Y_{1o}) and parametric asymmetry (Δm_{23}). When the structure is completely symmetric (i.e. $\Delta m_{23}=0$), the two loci intersect with each other at a frequency of 153 Hz and Y_{1o}=6.6%. Graphically (Fig 2-6a) as well as qualitatively, this is mode crossing. When the structure is asymmetric (e.g. $\Delta m_{23} > 0$), the two loci do not intersect. Instead, they approach each other and then diverge as Y_{1o} increases. Thus the point when Y_{1o}=6.6% is the veering point.

In the 3rd and 4th modes, the Y direction vibrations of the two secondary stages (i.e. y_2 and y_3) are dominant in the respective eigenvectors (see Table 2-3). Therefore, only the evolution of y_2 and y_3 are plotted in Fig 2-6. Fig 2-6b shows such evolution of the 3rd mode. When $\Delta m_{23} = 0$ and

Y_{1o} = 0, y_2 is 0.7 and y_3 is – 0.7 (i.e. opposite directions). When Y_{1o} reaches the crossing point, the value of y_3 suddenly changes from -0.7 to 0.7 and stays constant with y_2 unchanged (i.e. y_2 and y_3 become in the same direction); in other words, modes 3 and 4 'swap' at the crossing point. This transition is more gradual (i.e. without a discontinuity) for non-zero Δm_{23}. Similarly, as shown in Fig 2-6c, in the 4th mode, y_2 and y_3 change from being in the same direction to opposite directions. Note that modes shapes shown in Fig 2-3 correspond to those before curve veering has happened.

To summarize, the eigenvalues and eigenvectors of the 3rd and 4th modes exhibit veering. Furthermore, for the asymmetric case ($\Delta m_{23} > 0$) at the veering point, y_3 becomes zero while y_2 becomes 1 in the 3rd mode, while opposite happens in the 4th mode. This is recognized to be mode localization [56].

The above analysis shows that curve veering exists in the SR4DoF flexure mechanism dynamics. However, the more important observation here is that the curve veering happens at about the same operating point (Y_{1o}=6.6%), positive parametric asymmetry ($\Delta m_{23} > 0$), and frequency (153 Hz) at which CNMP zeros appear (see Fig 2-4 and Fig 2-5). Moreover, the CNMP phenomenon and curve veering share the same key factors such as closely spaced modes, parametric asymmetry, and mode coupling (caused by operating point variation in this case). Therefore, we hypothesize that this is not merely a coincidence and that there exists a fundamental relationship between these two phenomena. If established, this would allow a new physical perspective and interpretation of the CNMP phenomenon, and may help guide physical system design.

2.4. Contributions and Conclusion

The key contributions of this chapter are: (1) a lumped-parameter modeling approach is proposed to analytically model the dynamics of flexure mechanisms comprising the parallelogram or double parallelogram modules. This model captures the key relevant geometric non-linearity in large displacement flexure mechanics. Linearization about any arbitrary operating point enables frequency domain analysis; (2) Based on this model we are able to

predict previously unexplained CNMP zeros seen experimentally. The model establishes the existence of CNMP zeros under certain combinations of operating point and parametric asymmetry in the non-collocated transfer function of a simple representative XY flexure mechanism; (3) This finding helps generate the design insight that, rather than an intuitively symmetric design, an intentional positive/negative asymmetry in mass can guarantee/avoid CNMP zeros, respectively. (4) This overall investigation may have relevance not just to the XY flexure mechanisms considered in this chapter but also to a broader range of flexible systems.

In addition, there are several questions posed by this chapter: (1) Experimental validation of the analytical predictions such as the CNMP map (Fig 2-5) that can validate the modeling simplifications and assumptions. (2) While CNMP zeros were predicted for the SR4DoF as well as the full XY flexure mechanisms, these results are mathematical; greater physical insight into what causes the CNMP zeros is desirable. (3) The potential correlation between CNMP zeros and curve veering was based on observations in this chapter but needs to be investigated scientifically. (4) Based on these findings, physical and control system design can be explored to achieve the originally stated dynamic performance goals of large range, high precision, and high speed. These questions are answered in the following chapters.

Chapter 3. On the Non-Minimum Phase Zeros of Lightly Damped Flexible Systems

In Chapter 2, an approach is proposed to model the relevant CNMP zeros and predict the conditions where these CNMP zeros appear. In this chapter, a more fundamental question is considered: what is the genesis and what is the physical meaning of CNMP zeros?

This chapter focuses on the zeros of systems that are linear time invariant (LTI) and single input single output (SISO). There are three main original contributions presented in this chapter. First, we present a systematic investigation using the variation of modal residues to demonstrate the conditions for the existence of various types of zeros in the lightly damped flexible systems. Fig 3-1 illustrates all types of zeros that can exist in such systems, namely marginal minimum phase (MMP), real minimum phase (RMP), real non-minimum phase (RNMP), complex minimum phase (CMP), and complex non-minimum phase (CNMP). The term marginal minimum phase (MMP) is based on the assumption of lightly damped flexible systems and refers to complex zeros that are slightly to the left of the imaginary axis due to the small damping. Given the significance of NMP zeros in control system design and performance, particular emphasis is provided on RNMP and CNMP zeros.

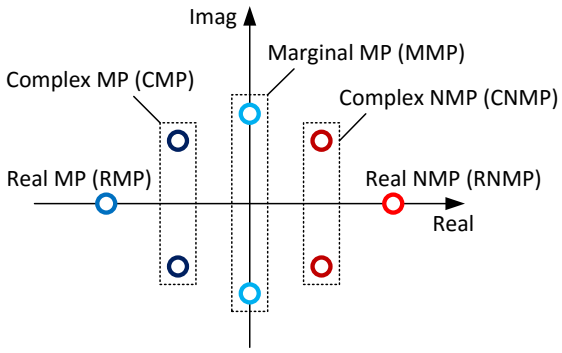

Fig 3-1 Definition of all types of zeros discussed in this chapter.

Second, we propose a graphical tool that utilizes the sequence of dominant modes and the sequence of their relative signs to predict the types of zeros of lightly damped flexible systems. Via these sequences, one can physically understand the sufficient and necessary conditions for the existence of certain types of zeros; one can also explain the initial transients of the time-domain step responses of systems with certain type of zeros. Moreover, this tool helps visualize the impact on close loop control systems with NMP zeros. Third, we show the capability of extending this investigation to cover more general flexible systems, and not just flexure mechanisms.

The rest of this chapter is organized as below. Section 3.1 provides a literature review on modal theory (i.e., modal decomposition and the variation of modal residues) and various kinds of zeros in flexible systems. Section 3.2 starts with the assumptions and definitions that are used in the rest of the chapter. Next, we introduce the concept of "sequence of dominant modes" to explain the conditions for the existence of various types of zeros. Particular emphasis is placed on NMP zeros, including necessary and sufficient conditions, physical interpretations, and impact on closed loop control. Additionally, we explain why flexible systems with different NMP zeros behave differently in the initial transient of their time-domain step responses (i.e., initial undershot, zero-crossings, and monotonousness). In Section 3.3, we explore a situation of flexible systems with both RNMP and CNMP zeros that have non-zero damping. Additionally, other generalizations are discussed. Lastly, this chapter concludes in Section 3.4.

3.1. Literature Review

In recent years, there has been a growing body of research literature on the zeros in the dynamics of flexure mechanisms that are LTI systems. These studies focus on properties of zeros in two domains: the frequency domain and the time domain. In the frequency domain, the studies can be further classified into two groups: (1) studies that focus on several fundamental system types; and (2) studies that focus on certain specific types of zeros based on various specific systems.

Studying the zeros of LTI systems using fundamental system types is based on the idea of modal decomposition [57]. Since a single mode cannot lead to any zeros, the simplest system type to study zeros is a system with only two modes. This representation can be obtained from specific systems with two degrees of freedom (DoF), or systems that can be truncated into only two dominant modes. In this system type, an even simpler case is when one of the modes has zero natural frequency (i.e., the rigid-body mode). Miu [58] used a two-mass torsional model and studied the variation of the locations of zeros (also termed as the migration of zeros) due to the variation of sensor combination. Rankers [59] studied the intersections between a rigid-body mode and a flexible mode in a Bode plot. He demonstrated that the variation of zeros due to the variation of modal residues. Coelingh [60] studied a motion stage with flexible guidance and shows the mapping between sensor / actuator locations and various types of zeros. In this body of works, the MMP zeros, RMP zero, and RNMP zero are identified under certain conditions.

There are also studies that focus on systems that are generalized from two-mode model. Tohyama and Lyon [48, 61] studied the transfer function in room acoustics and used a system with two modes and a constant remainder. By varying the modal residues and the remainder, they identified MMP zeros and CNMP zeros. However, these studies are limited by two main factors. First, the mathematical investigation requires further work to obtain a complete architecture, since these studies only provided the variation of the remainder without showing the relation between the two modes. As result, some type of zeros (e.g., RNMP zero) are not captured. Second, few physical insights are provided for the conditions that occur when zeros change their types. Duffour and Woodhouse [62] studied the transfer function of linearized systems with a frictional contact. In their studies, analytical expression and the locus technique are used to examine cases with two modes only, two modes with a constant remainder, and three modes, as well as undamped and damped systems. MMP zeros and CNMP zeros are identified; however, in each of the cases reported, attention is only given to a specific condition without a complete scenario (e.g., the RNMP zero can appear in a two-mode model but is not covered). Martin [43] studied a model with three modes numerically and identified the MMP, RNMP, RMP, and CNMP zeros. However, his results are limited to the numerical results without further generalization. It is noteworthy that he also considered the transfer function decomposed into a general form with n modes that are second order and undamped. He concluded that for the

situation of sensor and actuator collocation, the zeros are guaranteed to be MMP zeros and are located between the poles of the system (known as poles-zeros alternation).

The advantage of this group of work (using generalized forms of systems) is that when certain assumptions are met, the transfer function of a general system can be regarded as a combination of several of such simple forms. Thus the conclusions developed based on these forms can be applied. However, most of these studies simply report a mathematical or experimental observation, without providing further insight into when and why the zeros appear, especially RNMP and CNMP zeros.

The second group of studies focuses directly on a certain type of zeros based on specific systems. Miu [63] studied the MMP zeros in a simple class of systems in the form of single-axis serially connected spring-mass structures. He concluded that for this class of systems, the MMP zeros indicate the natural frequencies of several sub-systems isolated by the actuator and sensor. Chandrasekar et al. [64] showed that all of the zeros obtained in such systems with serially connected spring-mass structures are MMP zeros. Straete [65] used the approach of bond graph to study zeros generally and concluded that zeros are related to subsystems where the energy is "trapped". Also, Calafiore [66] analytically characterized the sub-systems related to the zeros. However, in these works, a compact sub-system is available only in a simple class of systems. For a general system (especially when the output is a linear combination of the states of the system), the related sub-systems are difficult to identify and their physical insights are not easy to obtain. An example is Coelingh's model [60] (Fig 3-7a), which uses a multi-axis system and has certain coupling between the modes in different axes.

Studying the zeros of systems with a single beam has also been an active area of research. Spector and Fleshner [42, 45] studied a non-collocated pinned-free beam model and identified the migration of zeros on the two axes due to variation of the sensor location. Wie and Bryson [67] studied the pole-zero patterns in flexible structures including beams, membranes, and triangular trusses. Lee and Speyer [68] used a Bernoulli-Euler beam and studied the migration of zeros in various input-output transfer functions. In addition, Aphale et al. [69] studied the zeros of a cantilever beam with the impact of a feed-through term, and Vakil et al. [70] studied the location of zeros of a single flexible beam under the variation of different physical parameters. In these works, the migration of zeros is shown to be anchored on the real and imaginary axes.

However, the migration of zeros can only show the conditions of RNMP zeros, which is not sufficient to obtain the physical interpretation of RNMP zeros. One particular study of a cantilever beam presents an attempt at establishing a connection between RNMP zeros and wave propagation [58]. However, this theory is limited by two main factors. First, the scope is limited to only continuous systems whose equation of motion contains a fourth order differentiator (e.g., Euler-Bernoulli beams). Therefore, this would not apply to the vibration of a string. However, RNMP zeros exist in a much broader range of systems, thus generalization is desired. Second, part of the derivation provided in this work is incorrect and requires a re-derivation.

Though CNMP zeros have been observed in flexible systems [16, 71], there are a limited number of studies that focus on understanding these zeros. Cannon and Schmitz [47] identified RNMP zeros and CNMP zeros numerically in the transfer function of a pinned-free beam. Loix et al. [49] studied a four DoF spring-mass model with spring stiffness variation. They identified the existence of CNMP zeros and the corresponding zero locus. They also provided an experimental validation on the existence of CNMP zeros in a cantilever beam set-up. It is interesting to note that the CNMP zeros captured in both works are not common in the theoretical studies of beams (as reviewed in the previous paragraph). However, few discussions are available to understand this difference. Hoagg et al. [50] investigated a three DoF spring-damper-mass model that also captured CNMP zeros. However, they assumed a large damping ratio ($\zeta>1.3$) to create CNMP zeros. In a physical system example, Awtar [51] predicted and observed CNMP zeros arising in the non-collocated transfer function of a multi-spring multi-mass servo system as a result of the electromagnetic coupling between a DC motor and tachometer. In our recent work, CNMP zeros have been experimentally observed and modeled in a flexure mechanism-based motion stage that is lightly damped [72]. However, all of these studies on CNMP zeros are case-specific and cannot be generalized to provide a further understanding on a broader range of flexible systems.

The advantage of this group of works that focuses on specific systems is that the existence of certain type of zeros (particularly the RNMP and CNMP zeros) can be validated via physical models and experiments. Furthermore, the relationship between physical parameters and the locations of zeros are demonstrated. However, there is lack of physical interpretation on the

conditions when zeros change to NMP zeros, especially CNMP zeros. Also, it is desirable to have a generalized conclusion that can unify all these studies.

There is another large body of studies that focus on the zeros in the time-domain response. Generally, zeros are understood as the source of a blocking effect, i.e., each zero blocks a specific input signal [16, 73]. It follows that for a system with NMP zeros, an exponential unbounded input determined by the system's NMP zeros leads to a bounded output. It is also commonly known that the step response of a stable, strictly proper transfer function exhibits initial undershoot if and only if the system has an odd number of positive zeros [73, 74], where the initial undershoot is defined for a step response whose initial value and steady state value have opposite signs. Great effort is made to understand the relationship between the number of NMP zeros and the time-domain properties of a system in response to a step input [75-77]. However, for the initial properties (e.g., undershoot, zero-crossings, and monotonousness), Hoagg and Berstein [73] reviewed the relevant studies and found the statements to be inconsistent with each other. This implies that the connection between NMP zeros and the properties of corresponding step responses is still open to consideration. Furthermore, to understand the initial response of a system to a step input, the information on the numbers of the system's NMP zeros alone is insufficient. In addition, in a physical example, Cannon and Schmitz [47] experimentally observed the initial undershoot of a cantilever beam with NMP zeros and explained that the time for the wave to propagate the length of the beam is non-zero. All of these works present a time-domain understanding of the NMP zeros. However, similar to the previously discussed groups of works, few physical insights are available on why the NMP zeros appear. Also, attention is mostly paid on RNMP zeros, while the understanding of CNMP zeros is lacking.

To summarize, the gap of knowledge in this area is characterized into three main points as follows. First, the technique of varying modal residues can be used to study general flexible systems. However, current results are incomplete to present analysis on systems that include all possible types of zeros. Second, regarding the NMP zeros (especially CNMP zeros), the interpretation of the physical meaning and the genesis (when and why) of NMP zeros is still an open question. Third, current available references are presented as isolated case studies; therefore a systematic presentation to demonstrate the connection between them is desired.

In order to address the above-mentioned knowledge gap, we present a systematic step by step investigation based on modal decomposition and the variation of modal residues in the next section. We begin with preliminaries with a set of assumptions that can help one better define the systems being considered. Several general system types are presented in a sequence that gradually increases complexity.

3.2. Investigation of Zeros based on the Variation of Modal Residues

3.2.1. Assumptions and Definitions

For an LTI and SISO system, the input-output dynamics (i.e., transfer function) can be expressed as the sum of the contribution of its decomposed modes (Eq. (3-1)), where all of the coefficients in both the numerator and the denominator are real numbers. It follows that the variation of modal residues (α_i) leads to variation of the coefficients of the numerator (b_i), and thus the variation of zeros of *G(s)*. In addition, the variation of p_i would change both a_i and b_i. However, since the variation of p_i refers to the change of the internal modes of the system. As the first step investigation, such variation is not considered in the main section of this chapter. We will discuss this situation in Chapter 4.

$$G(s)=\frac{b_m s^m+\cdots+b_1 s+b_0}{a_n s^n+\cdots+a_1 s+a_0}=\sum_{i=0}^{n}\frac{\alpha_i}{s+p_i} \tag{3-1}$$

Three main assumptions are made here. These assumptions are consistent with the investigations in the field of structural dynamics in the literature.

Assumption A3-1: the system is a structural system. In other words, all of the decomposed modes are second order, and there are no first order modes.

Assumption A3-2: damping is negligible.

Assumption A3-2 is valid for flexible systems such as flexure mechanisms that are monolithic with no rolling or sliding joints [2], as well as for space structures and machines that operate in vacuum [78]. Our experiments in flexure-based motion systems [72] confirmed that

the dissipation from material hysteresis and viscous effects is very low, with a damping ratio in the order of 0.001.

Assumption A3-3: the system considered has no right half plane poles (i.e., the plant is open loop stable).

With Assumptions A3-1 to A3-3, Eq. (3-1) can be restated as Eq. (3-2), where ω_i is the natural frequency of the ith mode.

$$G(s) = \sum_{i=0}^{k=n/2} \frac{\alpha_i}{s^2 + \omega_i^2} = \frac{N(s)}{D(s)} = \frac{\prod (s^2 + c_j)}{D(s)} \tag{3-2}$$

Note that in Eq. (3-2) the modal residue (α_i) is the residue related to the ith decomposed mode (i.e., a second order system). Based on the theory of modal decomposition [59], α_i is determined by the ith eigenvector, the system's mass matrix, and the sensor / actuator locations. In this chapter, we only consider α_i as real numbers.

With Eq. (3-2), it follows that the numerator of *G(s)* has no odd order terms; therefore, the zeros of *G(s)* are always in symmetric pairs (i.e., symmetric w.r.t. to either the two axes or the origin). Based on the value of c_j, there are three different situations. First, if c_j is positive real, then the zeros are located on the imaginary axis and symmetric w.r.t. the real axis. Second, if c_j is negative real, then the associated zeros are located on the real axis and symmetric w.r.t. the imaginary axis. Among these two zeros, one is an RNMP zero, and the other is a RMP zero. Third, if c_j is complex, there must be a c_j^* as its complex conjugate. Thus the associated zeros are in the form of a quartet in the complex plane, and they are symmetric w.r.t. to both the real and the imaginary axes. Among these four zeros, there is a pair of CNMP zeros and a pair of CMP zeros.

For easy representation, in the naming convention of this thesis, we use the term marginally minimum phase (MMP) zeros to refer to the first situation (a pair of zeros on the imaginary axis, see Fig 3-1). This because even though we have assumed zero damping, in practice there is a small amount of viscous or structural damping, which shifts the zeros slightly to the left of the imaginary axis, making them marginally minimum phase. Next, we use the term RMP-RNMP zeros to refer to the zeros in the second situation (a pair of zeros with one RMP zero and one

RNMP zero). Lastly, we use CMP-CNMP zeros to refer to the zeros in the third situation (a quartet of zeros with a pair of complex MP zeros and a pair of CNMP zeros).

It is noteworthy that based on Assumptions A3-1 to A3-3, all of the NMP zeros are in the form of either RMP-RNMP or CMP-CNMP zeros. Furthermore, MMP zeros and RMP-RNMP zeros have only one pair of zeros, while CMP-CNMP has two pairs (i.e., a quartet).

In the remainder of this section, several new terms are defined to facilitate the subsequent discussion. The purpose of introducing these definitions is described below: First, we setup the concept of mode "take over"; Second, the idea mode "take over" is used to define the "dominant mode" in a specific frequency region; Finally, the sequence of the dominant modes from low frequency to high frequency is used as the proposed tool to examine the various types of zeros identified above.

Note that under the Assumptions A3-1 and A3-2, the phases of a transfer function can only be either 0° (defined as the positive direction / sign) or 180° (defined as the negative direction / sign). For a system's transfer function $G(s)$, denote the contribution of mode x as $G_x(s)$ and denote the contributions of all other modes as $G_{\bar{x}}(s)$. It follows that $G(s)=G_x(s)+G_{\bar{x}}(s)$.Then consider the following situations:

(1) If for a mode x, there is a frequency ω_0 such that
    a. $G(j\omega_0)=0$, and
    b. $G(j\omega_0^+)\cdot G_x(j\omega_0)>0$ or $G(j\omega_0^-)\cdot G_x(j\omega_0)<0$

    Then we say that mode x *takes over* at ω_0, as frequency increases. Note that condition (1a) indicates that there is zero crossing in $G(s)$ at ω_0, and condition (1b) indicates that this zero crossing is caused by mode x.

(2) If there are multiple modes (e.g., mode x_1, mode x_2, and mode x_3) satisfying the above condition (1), then rewrite $G(s)$ into the form as $G(s) = G_{x1}(s) + G_{x2}(s) + G_{x3}(s) + G_{rest}(s)$ and consider the following.
    a. Consider the sole impact of each candidate. For example, $G_1(s)=G_{x1}(s)+G_{rest}(s)$. If for G_1, check if there is a frequency ω_1 sufficient closed to ω_0, such that condition (1) is satisfied.
    b. If there is at least one candidate (e.g., mode x_1) that satisfies condition (2a), then we say that mode x_1 takes over at ω_0, and those that do not satisfy condition (2a)

are not the modes that take over at ω_0.

c. If there are multiple candidates that satisfy condition (2a), then we can use an abstract mode to represent the contribution of these candidates. In this case, the only possible cases are: DC constant for low frequencies and double integrator (DI) for high frequencies.

d. If none of the candidate satisfies condition (2a) but recall that each of them satisfies condition (1), then we say there is an abstract mode that takes over at ω_0, as the contributions of all the candidates.

(3) If for a mode x, there is a frequency ω_0 such that

a. $|G_x(j\omega_0)| >> |G_{\bar{x}}(j\omega_0)|$, and

b. Among all of the decomposed modes, there is at least one mode that is non-decreasing.

Then we say that mode x *takes over* at ω_0, as frequency increases. Note that this is a take-over where there is no zero-crossing in $G(s)$. In addition, condition (3b) indicates that there is no take-over defined in high-frequencies region where all of the decomposed modes are decreasing.

(4) Mode x cannot take over twice continuously. In other words, if mode x has two take-overs (or more), then between each of the two take-overs of mode x, there has to be a take-over of a mode other than mode x.

(5) We say that mode x is the *dominant mode* in the range (ω_1, ω_2). If:

a. Mode x takes over at ω_1, and mode y takes over at ω_2, with $\omega_1 < \omega_2$.

b. There is no mode that takes over in the range (ω_1, ω_2).

Particularly, if there is no mode that takes over in the range (0, ω_1), then mode x is the dominant mode in the range (0, ω_2), or the dominant mode in the low frequencies; if there is no mode that takes over in the range (ω_2, $+\infty$), then mode y is the dominant mode in the range (ω_2, $+\infty$), or the dominant modes in the high frequencies. Also, we say that mode x is *taken over* by mode y at ω_2.

With this definition, it follows that at certain frequency range, a system behaves as the dominant mode corresponding to that frequency range. Moreover, the phase of a system in that frequency range is determined by the dominant mode.

Note that the definition above has some rough definitions (e.g., ">>" and "near") that result in rough bounds. However, we are interested in the sequence of dominant modes of a system. As illustrated later in this chapter, such sequence along the entire frequency axis is well-defined even with the rough positions of the bounds. Moreover, such sequence has significant physical meaning related to the NMP zeros. Without losing generality, in this thesis, these sequences are discussed in the direction as frequency increases.

3.2.2. Systems with Two Modes: a Double Integrator and a Second Order Mode

Based on Assumptions A3-1 to A3-3, consider the assumptions as following.

Assumption A3-4: consider the transfer function $G(s)$ of a system and its two modes of interest (denoted as mode u and mode v). In the frequency range between the natural frequencies of mode u and mode v, $G(s)$ can be approximated as the contributions of only these two modes, while the contributions of all other modes are negligible.

Assumption A3-5: mode u is much slower than mode v (i.e., $\omega_u << \omega_v$).

Therefore, with assumptions A3-1 to A3-5, in the frequency region around ω_v, mode u can be approximated as a rigid-body mode (double integrator). The first system type is then presented. Consider the variation of modal residues of only two modes as in Eq. (3-3), where mode u is a double integrator.

$$G(s) = \frac{\alpha_u}{s^2} + \frac{\alpha_v}{s^2 + \omega_v^2};\ \omega_v > 0;\ \alpha_u > 0 \tag{3-3}$$

Additionally, for ease of representation, we follow the assumption as below, which explains why α_u is positive in Eq. (3-3).

Assumption A3-6: for the transfer function of a system that is expressed as the contributions of its relevant decomposed modes, the slowest mode is in the positive direction (i.e., the modal residue of the slowest mode is positive).

With Eq. (3-3), if we vary α_v, but fix every other variable, then we obtain the locus of zeros along the real and the imaginary axes as shown in Fig 3-2a, which is termed as *the migration of zeros* in previous studies [58, 60, 68]. Fig 3-2b demonstrates the corresponding values of α_v at certain zero locations shown in Fig 3-2a. Note that Fig 3-2a can also be obtained via the locus technique [79]. Therefore, Fig 3-2a is also termed as *the locus of zeros*.

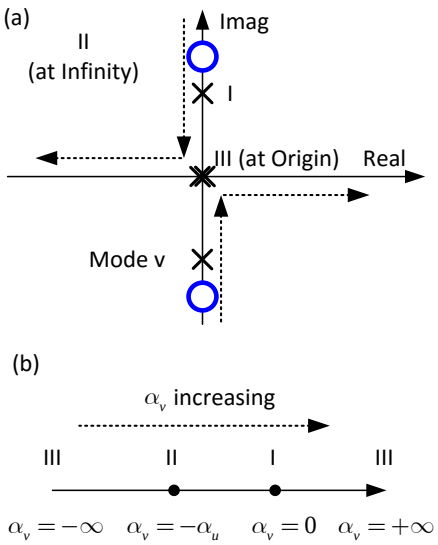

Fig 3-2 Migration of zeros (double integrator + single mode). (a) the locus of zeros; (b) conditions of α_v.

Three main observations are reviewed. First, when α_v is in the segments of I-II and I-III in Fig 3-2b, the corresponding zeros appear in the form of MMP zeros as in Fig 3-2a. Particularly, when α_v is at I or III in Fig 3-2b, the zeros are collocated with the poles of mode v or mode u (at origin) respectively, leading to a pole-zero cancellation. Note that in Fig 3-2b, the situations of both positive infinity and negative infinity are related to same point III. Second, when α_v is in the segment of II-III in Fig 3-2b, the zeros appear in the form of RMP-RNMP zeros. Moreover, RMP-RNMP zeros are the only type of NMP zeros in systems described as Eq. (3-3). Third, for systems with collocated sensor and actuator, one can guarantee $\alpha_u \cdot \alpha_v > 0$ (i.e., $\alpha_v > 0$ under Assumption A3-6) [59]. Consequently, the zeros are guaranteed to be MMP zeros and appear between the poles (known as pole-zero alternation) [43]. However, for systems with non-collocated sensor and actuator, there is no such guarantee. Instead, zeros can appear in any situations shown in Fig 3-2.

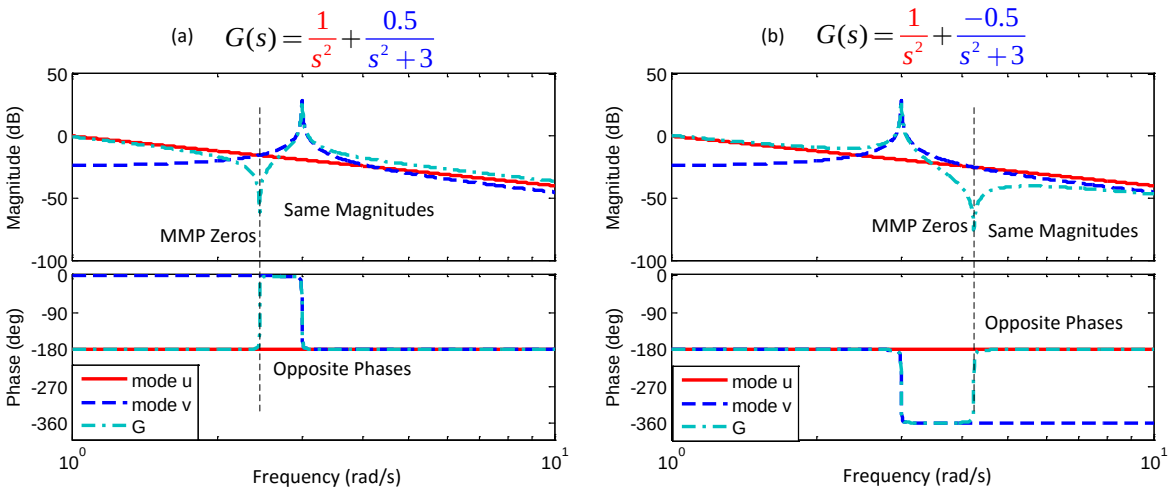

Fig 3-3 Bode plot of *G(s)* expressed in decomposed modes showing MMP zeros: (a) decomposed modes are in-phase; (b) decomposed modes are out-of-phase.

To demonstrate how MMP zeros appear, a numerical examples of two fundamental situations are presented in Fig 3-3 (similar works can be found in [57, 58]). In both situations, the transfer function *G(s)* has only two modes, namely mode u (the rigid-body mode) and mode v. The Bode plots of *G(s)* are then examined. With Assumptions A3-1 and A3-2, a transfer function can only be either positive or negative at a given frequency. Therefore, in Fig 3-3a, the modes with the same sign (i.e., $\alpha_u \cdot \alpha_v > 0$) are called *in-phase* modes; conversely, in Fig 3-3b, the modes with opposite signs (i.e., $\alpha_u \cdot \alpha_v < 0$) are called *out-of-phase* modes. In each plot of Fig 3-3, there is a specific frequency when the two decomposed modes have the same magnitudes but different signs (i.e., opposite phases). At these frequencies, $G(s) \rightarrow 0$, indicating a pair of zeros on the imaginary axis. To conclude, for system with only two modes and under Assumptions A3-1 to A3-6, the MMP zeros appear if and only if (sufficient and necessary condition) there is a frequency where these modes have the same magnitude but opposite signs.

Since MMP zeros are extensively studied, we will focus on the NMP zeros in the rest of this chapter. Furthermore, since in-phase modes are guaranteed with only MMP zeros, we will focus on the situation of out-of-phase modes (i.e., $\alpha_u > 0$ and $\alpha_v < 0$ under Assumption A3-6).

Shown in Fig 3-4 is the Bode plot of a system that described as Eq. (3-3) which contains a double integrator (mode u) and a second order mode v. The comparison between the MMP case and the RMP-RNMP case is demonstrated, under the variation of the modal residue α_v ($\alpha_v < 0$).

Under the Assumptions A3-1 to A3-3, only the magnitudes are plotted, with the signs of the magnitudes indicating their corresponding phases.

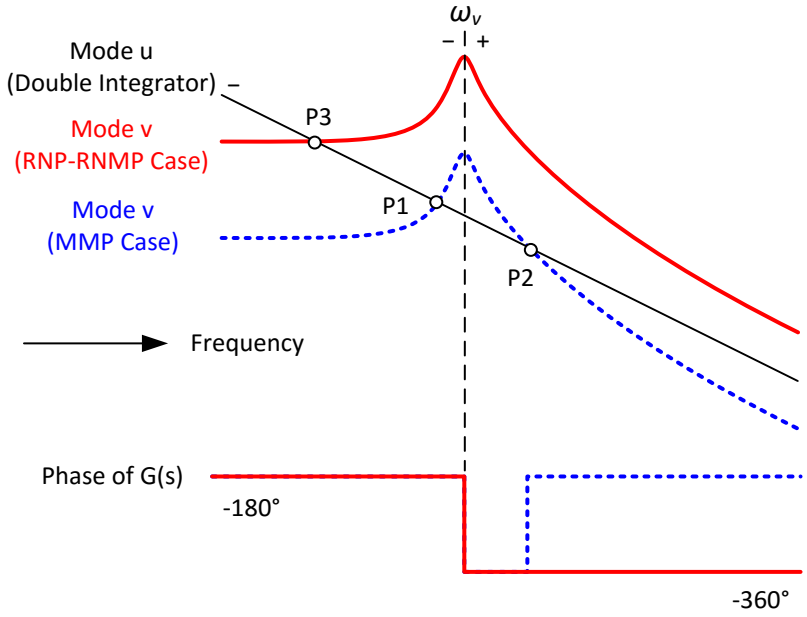

Fig 3-4 Decomposed modes under different modal residues (double integrator + single 2nd order mode).

The blue dashed line corresponds to mode v in the case of MMP zeros. Two points (P1 and P2) indicate the same magnitudes of the two modes. At P1, the contributions of the two modes are in the same direction (negative). According to the definition of mode "taking over", there exist a frequency near P1 where mode u is taken over by mode v (i.e., a take-over in the same direction). In contrast, the contributions of the two modes are in the opposite directions at P2. Thus, P2 is the frequency where mode v is taken back over by mode u (i.e., a take-over in opposite directions). Therefore, the sequence of dominant mode in this MMP case is denoted as [u]-[v]-[u]. Moreover, P2 indicates the existence of MMP zeros (same magnitude but opposite directions). Thus, P2 is the frequency of the associated MMP zeros. Conversely, the red solid line corresponds to mode v in the case of RMP-RNMP zeros. In this case however, there is only one point (P3) indicating that there is a taking over from mode u to mode v near P3. Consequently, the sequence of dominant modes is denoted as [u]-[v].

Next, the phase of *G(s)* is examined. According to Eq. (3-3), the phase of *G(s)* will follow the phase of the mode that is the dominant mode. Thus, for the MMP case, at low frequencies,

G(s) starts as the double integrator whose phase is -180°. Next, near P1, mode v takes over and becomes the dominant mode. Thus, *G(s)* drops towards -360° due to the resonance of mode v. However, at P2, the double integrator takes over again. Consequently, *G(s)* rises back to -180°. In contrast, there is only one mode taking over in the RMP-RNMP case. Therefore, near P3, mode v takes over and is never taken back over by the double integrator. As a result, *G(s)* begins at -180° and drops to -360° due to the resonance of mode v.

A numerical example is provided to show a comparison between an MMP case and an RMP-RNMP case. The system used in this example is shown in Eq. (3-4). We use α_v=-0.5 for the MMP case and α_v=-5 for the RMP-RNMP case. Additionally, a small damping ratio of 0.001 is added to avoid numerical errors. Fig 3-5a is the Bode plot of the MMP case with Fig 3-5c as its corresponding step response. As a comparison, Fig 3-5b is the Bode plot of the RMP-RNMP case with Fig 3-5d as its corresponding step response.

$$G(s) = \frac{1}{s^2 + 0.002s + 1^2} + \frac{\alpha_v}{s^2 + 0.02s + 10^2} \tag{3-4}$$

Three main observations are presented. First, for both Bode plots (Fig 3-5a and Fig 3-5b), since mode u is much slower than mode v, mode u is then approximated as a double integrator around the natural frequency of mode v. Second, in the MMP case (Fig 3-5a), since the sequence of dominant modes is [u]-[v]-[u], the phase of *G(s)* eventually rises back to -180°. In the RMP-RNMP case (Fig 3-5b) however, since the corresponding sequence is [u]-[v], the phase of *G(s)* eventually drops to -360°. Third, there is an initial undershoot in step response of the RMP-RNMP case (Fig 3-5d), while no such observation is presence in the MMP case (Fig 3-5c).

The step response can be explained by the sequence of dominant modes. For the MMP case, since the sequence is [u]-[v]-[u] (Fig 3-5a), one can see that mode v only dominates a relative small frequency region, while mode u dominates in most of the frequency regions. Consequently, the step response of *G(s)* is essentially the step response of mode u, which is a second order system in the positive direction (Fig 3-5c).

As a comparison, for the RMP-RNMP case, since the sequence is [u]-[v] (Fig 3-5b), one can see that mode v is the dominant mode throughout the entire high frequency region. High frequency properties are related to fast transient behavior. This is evident in the step response of

G(s), where the initial transient behavior is dominated by mode v (i.e., a second order system in the negative direction). Therefore, in Fig 3-5d, *G(s)* is initially dominated by mode v, resulting in an initially "wrong" direction. As time progresses, the slower mode (mode u) dominates, thus *G(s)* is eventually driven back by mode u to the positive direction (the "right" direction).

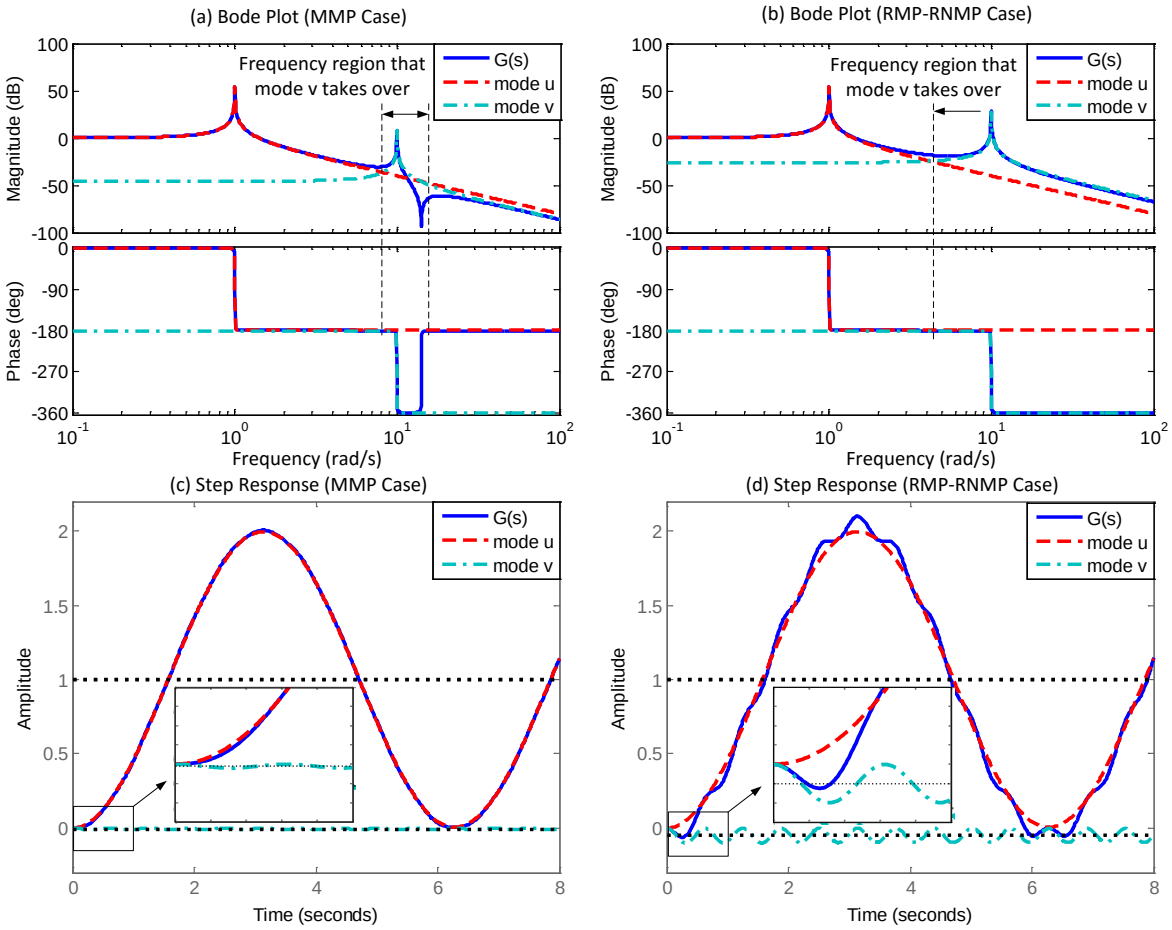

Fig 3-5 Numerical comparison: MMP case vs. RMP-RNMP case. (a) Bode plot of the MMP case; (b) Bode plot of the RMP-RNMP case; (c) step response of the MMP case; (d) step response of the RMP-RNMP case.

Based on the above discussion, one can see that there is a mapping from sequences to the types of zeros. Moreover, this mapping is bijective, under Assumptions A3-1 to A3-6. The reason is that, the sequence of dominant modes is a graphical representation of the analytical conditions of modal residues that lead to different type of zeros.

To further understand the physical insights of these sequences (particularly the RMP-RNMP case), Fig 3-6 illustrates the Nyquist plots of *G(s)* for both the MMP case and the RMP-RNMP

case. For ease of presentation, only the positive frequencies of the Nyquist curve are shown. The directions of the Nyquist curves indicate the increase of frequencies.

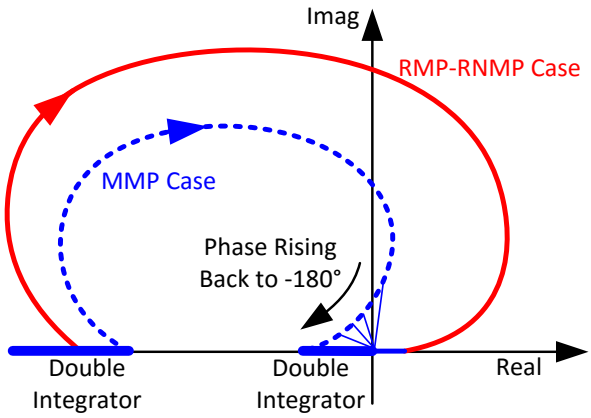

Fig 3-6 Nyquist plot of *G(s)*: MMP case vs. RMP-RNMP case.

In Fig 3-6, the double integrator (mode u) is a straight line on the left side of the real axis (always -180°). In contrast, mode v (i.e., a second order system in the negative direction) is associated with the trajectory in the shape of a circle. This circle grows to the left side along the positive direction of the Nyquist curve. Additionally, such a circle is clockwise and thus implies phase dropping as the frequency increases.

For the MMP case, since the sequence is [u]-[v]-[u], the double integrator (mode u) is the dominant mode at high frequencies, the circle ends on the left side of the origin. Therefore, if evaluating a vector from the origin to the circle, one can see that the phase of *G(s)* first drops due to the clockwise circle and then rises back to -180°. In the RMP-RNMP case however, since the sequence is [u]-[v], mode v is the dominant mode along the high frequencies, the circle ends on the right side of the origin. Thus, the phase of *G(s)* drops from -180° to -360°.

Therefore, for systems described as Eq. (3-3), RMP-RNMP zeros appear if the following two conditions are met. First, there exists a negative direction decomposed second order mode (i.e., mode u and mode v are out-of-phase). Second, the two modes subscribe to a certain sequence of dominant modes denoted as [u]-[v] (i.e., each mode dominates at its corresponding frequency region).

To illustrate a physical example of a negative direction decomposed second order mode, consider a cart that can rotate w.r.t. its center of mass as shown in Fig 3-7a [60, 80]. The transfer function T from the actuator force to the sensor displacement is examined. Based on the sensor and actuator configuration, there are two modes of interest: (1) the rigid-body mode; and (2) the rotational mode associated with the two vertical springs. In this manner, the rigid-body mode and the rotational mode refers to mode u (double integrator) and mode v discussed in this section, respectively. Regarding the rotational mode, if the actuator and the sensor are on opposite sides of the center of mass, then it becomes a negative direction decomposed second order mode of T. Moreover, if the actuator and the sensor locations meet certain conditions [60, 80] such that the rotational mode is the dominant mode throughout the high frequency region, then RMP-RNMP zeros are expected in T.

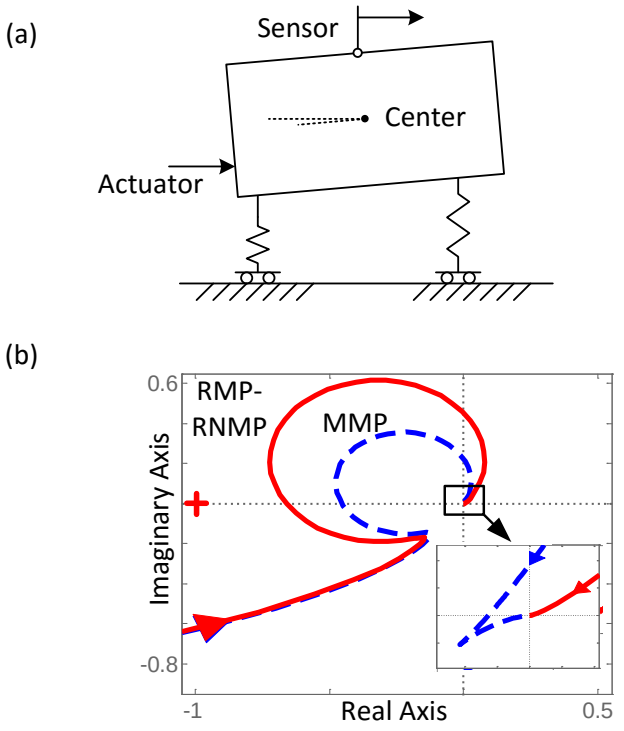

Fig 3-7 Example of second order decomposed modes in the negative direction and its impact on closed loop control (a) Rotatable spring-mass cart; (b) Nyquist plot of open loop transfer function.

This negative direction second order mode is important, because a problem is encountered when one attempts to apply a closed loop control while pursuing a high bandwidth. We use the

example in Fig 3-7a to illustrate this problem. Consider applying a classic closed loop control with unity feedback, where the crossover frequency is set to be between the natural frequencies of the translational mode and the rotational mode (usually with a lead filter). The Nyquist plot of the typical open loop transfer function L is shown in Fig 3-7b. The bulky circle is associated with the negative direction second order mode. Since this circle is on the left side of the Nyquist curve of L, it dramatically reduces the distance from the Nyquist curve to the critical point (-1, 0). This implies a peak in the sensitivity function $S=1/(1+L)$, which adversely affects the closed loop stability. Moreover, the size of the circle will significantly increase if the damping is decreased, making the stability problem more difficult. Additionally, such circles are presented due to the negative direction second order mode, regardless of whether it is an RMP-RNMP case or an MMP case.

Therefore, two points are noteworthy from the design standpoint. First of all, we recommend avoiding the negative direction second order mode via mechatronic design. In the example shown in Fig 3-7a, by simply placing the actuator and sensor on the same side of the center of mass, the rotational mode will change to be in a positive direction. This is consistent with previous literature [60, 80]. Secondly, if the crossover frequency is separated from the negative direction second order mode with an appropriate margin, one can apply a certain amount of phase lag (usually with a low pass filter) to “rotate” the problematic circle clockwise to a safety angular position [59, 81]. However, the crossover frequency requires careful consideration so that this additional phase lag minimally impacts the phase margin.

To conclude the discussion in this section, for systems described as Eq. (3-3), the relations between the sequences of the dominant modes, the corresponding segments in the locus of zeros (shown in Fig 3-2), and the corresponding type of zeros are summarized as Table 3-1. The situation of in-phase is also included such that the all situations in Fig 3-2 are covered. Note that we use the symbol [uv] to denote any possible sequence formed by mode u and mode v.

Table 3-1. The sequences of dominant modes and related zero types (double integrator + single mode).

Signs-Sequence	Sequence of dominant modes	Zero locus	Zero types	MMP Location
$\alpha_u>0$, $\alpha_v>0$ (in-phase)	[uv]	I-III	MMP	$(0, \omega_v)$ (pole-zero alternating)
$\alpha_u>0$, $\alpha_v<0$ (out-of-phase)	[u]-[v]	II-III	RMP-RNMP	N/A
	[u]-[v]-[u]	I-II	MMP	(ω_v, ∞)

In addition, using the decomposed view, one can also explain why a certain unbounded input (i.e., an exponential input related to the RNMP zero of a system) can lead to a bounded output. This is due to the fact that in the responses of the two decomposed modes, the unbounded parts have the same magnitudes but opposite signs, thus canceling out each other. However, this explanation cannot be visualized via the sequence of dominant modes. This is because the tool of the sequence can easily reveal the type of zeros but not the exact location of zeros.

3.2.3. Systems with Two Modes: Two Second Order Modes

A more general situation of systems with only two modes is achieved by removing Assumption A3-5. Consider the variation of modal residues of only two modes as in Eq. (3-5). Without losing generality, we assume that mode u is slower than mode v.

$$G(s) = \frac{\alpha_u}{s^2 + \omega_u^2} + \frac{\alpha_v}{s^2 + \omega_v^2};\ \omega_v > \omega_u > 0;\ \alpha_u > 0 \tag{3-5}$$

Since the analysis on systems described as Eq. (3-5) is very similar to the discussion in previous section, we simply present the results without detailed discussion. The locus of zeros is shown in Fig 3-8, where we vary α_v, but fix every other variable.

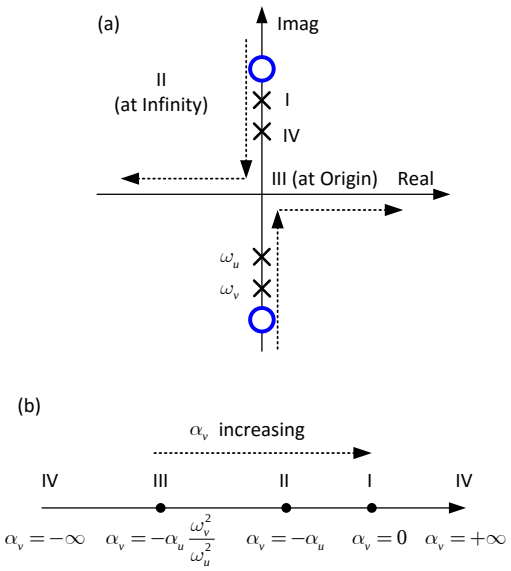

Fig 3-8 Migration of zeros (two 2nd modes): (a) the locus of zeros; (b) conditions of α_v.

The corresponding sequences of dominant modes are listed in Table 3-2. Note that comparing to Table 3-1, there is an extra sequence mapping to the segments III-IV in Fig 3-8. This is because mode u has a non-zero natural frequency.

Table 3-2. The sequences of dominant modes and related zero types (two 2nd order modes).

Signs-Sequence	Sequence of dominant modes	Zero locus	Zero types	MMP Location
α_u>0, α_v>0 (in-phase)	[uv]	I-IV	MMP	(ω_u, ω_v) (pole-zero alternating)
α_u>0, α_v<0 (out-of-phase)	[v]-[u]-[v]	III-IV	MMP	N/A
	[u]-[v]	II-III	RMP-RNMP	(ω_v, ∞)
	[u]-[v]-[u]	I-II	MMP	(0, ω_u)

3.2.4. Systems with Two Modes and a Double Integrator

In Section 3.2.2 and Section 3.2.3 we have covered the systems with only two modes. In this section, a more general class of systems is considered. Such systems have two modes and a remainder, following the assumption shown as below.

Assumption A3-7: consider the transfer function $G(s)$ of a system and the two modes of interest (denoted as mode u and mode v). In the frequency range near (ω_u, ω_v), the contributions of modes other than mode u and mode v are dominated by a single mode. Moreover, this mode is well-separated from the mode u and mode v.

In this thesis, the two modes of interest satisfying Assumption A3-7 are called the *closely spaced modes*. Following this definition, the two modes discussed in Section 3.2.2 and Section 3.2.3 are also called closely spaced modes. The difference between Eq. (3-5) and Eq. (3-6) is that Eq. (3-5) requires a much stronger separation between the closely spaced modes and the other modes. With Assumptions A3-1, A3-2, A3-3, and A3-7, consider systems described as Eq. (3-6), where *R(s)* is the contribution of other modes (*R* is termed as the remainder [48]). By Assumption A3-6, *R(s)* is dominated by a single mode, which is denoted as mode R in this chapter. In the analysis of this section, the resonance frequencies of mode u and mode v are fixed. Again, mode u is assumed to be slower than mode v.

$$G(s) = \frac{\alpha_u}{s^2 + \omega_u^2} + \frac{\alpha_v}{s^2 + \omega_v^2} + R(s);\ \omega_v > \omega_u > 0 \tag{3-6}$$

If mode R is much slower than mode u and mode v, then *R(s)* can be approximated as a double integrator; conversely, if mode R is much faster, then *R(s)* can be approximated as a constant (i.e., modal truncation). For flexure mechanisms, the first primary mode is usually the "rigid-body" mode at a much lower frequency. Therefore, *R(s)* is considered as a double integrator in this section. The situation of a fast mode R is discussed in the next section. As a result, Eq. (3-6) is restated as Eq. (3-7). Note that α_R is assumed to be positive by Assumption A3-6.

$$G(s)=\frac{\alpha_R}{s^2}+\frac{\alpha_u}{s^2+\omega_u^2}+\frac{\alpha_v}{s^2+\omega_v^2};\ \omega_v>\omega_u>0;\ \alpha_R>0 \tag{3-7}$$

Next, rewrite Eq. (3-7) into Eq. (3-8).

$$G(s)=\frac{\alpha_v s^2[(\alpha_u/\alpha_v+1)s^2+(\alpha_u\omega_v^2/\alpha_v+\omega_u^2)]+[\alpha_R(s^2+\omega_u^2)(s^2+\omega_v^2)]}{s^2(s^2+\omega_u^2)(s^2+\omega_v^2)} \tag{3-8}$$
$$=\frac{N(s)}{D(s)}\overset{\Delta}{=}\frac{\alpha_v\cdot A(s)+B(s)}{D(s)}$$

If a transfer function *T(s)* is defined as *T(s)=A(s)/B(s)*, one can state the following. First, the poles of *T(s)* are the poles of mode u and mode v. Second, *T(s)* has two pairs of zeros. One pair is fixed at the origin; the other pair follows a similar scenario discussed in Section 3.2.3 (two modes only), and is determined by the ratio α_u/α_v. Lastly, if we vary α_v while fixing other variables, then we obtain the root locus of *T(s)* with unity feedback, which is also the zero locus of *G(s)*. Also, if the sign of α_v is flipped, one obtains the complementary root locus.

The root locus and the complementary root locus of *T(s)* are shown as Fig 3-9. For ease of representation, only the first quadrant is shown in each case. Additionally, when α_u/α_v is fixed, the break in/away points can be computed by $d\alpha_v/ds=0$.

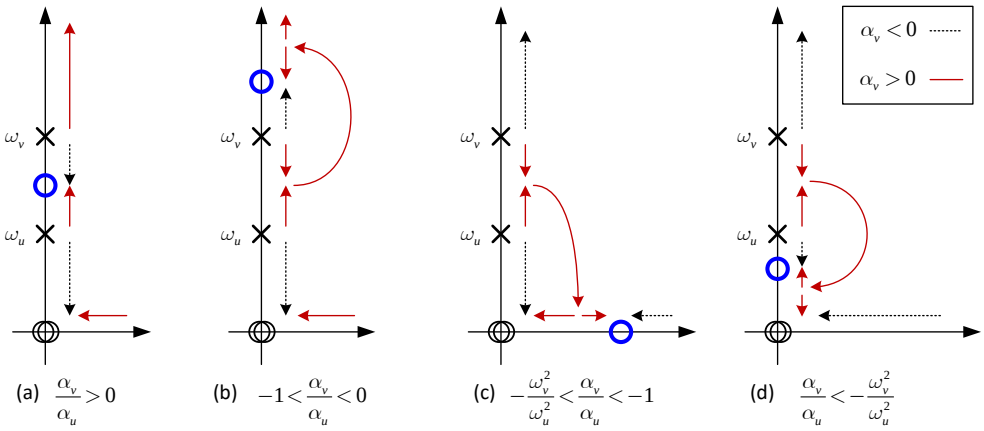

Fig 3-9 The locus of zeros under different values of α_u/α_v.

As the root locus of *T(s)* is the zero locus of *G(s)*, if the ratio α_u/α_v, is varied from -∞ to +∞, one obtains the family of the zero locus of *G(s)*. Fig 3-10 demonstrates a numerical example of zero locus family. There are some two cases where the loci are determined by the two natural frequencies ω_u and ω_v. Moreover, one can see that the loci in Fig 3-10 are no longer limited on the two axes (as Fig 3-8); instead, they fully cover the complex plane.

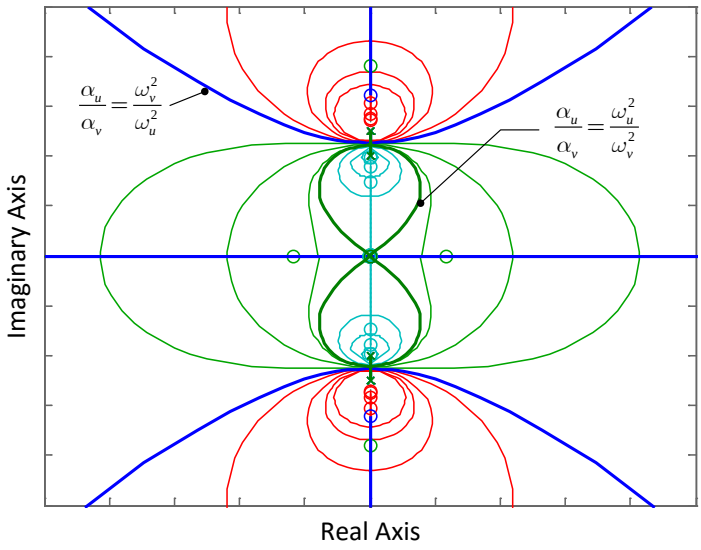

Fig 3-10 The family of the zero locus of *G(s)*.

From Fig 3-9, we can see the CMP-CNMP zeros appear only when $\alpha_u<0$, and $\alpha_v>0$. For ease of illustration, we consider the case when $\alpha_R=1$, $\alpha_u=-\alpha_0<0$, and $\alpha_v=\alpha_0$. Consequently, the variation of α_0 leads to the locus of zeros referring to the red solid locus shown in Fig 3-9c. As for the break in/away points, the associated α_0 is computed as $\alpha_{01}=(\omega_u-\omega_v)/(\omega_v+\omega_u)$ at the imaginary axis and $\alpha_{02}=(\omega_u+\omega_v)/(\omega_v-\omega_u)$ at the real axis.

Fig 3-11 demonstrates a Bode plot with different values of α_0. Since the double integrator has $\alpha_R>0$, mode R always exhibits a negative direction (-180°). Additionally, since $\alpha_v>0$ and $\alpha_v<0$, mode u has a negative direction while mode v has a positive direction.

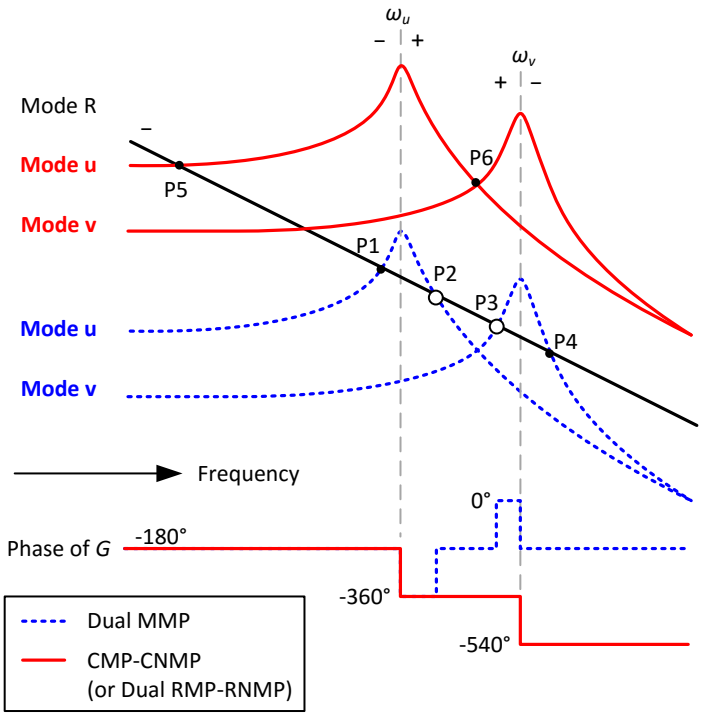

Fig 3-11 Decomposed modes under different modal residues (double integrator + two modes).

The blue dashed lines refer to the situations of mode u and mode v when $\alpha_0<\alpha_{01}$. Four points (P1, P2, P3, and P4) indicate the frequencies where mode R intersects with mode u and mode v (same magnitudes). Following a similar procedure as Section 3.2, one can see that there are take-overs near P1, P2, and P3. Moreover, P2 and P3 indicate the existence of two pairs of MMP (dual MMP) zeros nearby. However, there is no take-over near P4 according to the definition. Consequently, the sequence of dominant mode is [R]-[u]-[R]-[v].

The red solid line corresponds to the case when $\alpha_0>\alpha_{01}$. At low frequencies, P5 is similar to P1. At medium frequencies however, one can see that P2 and P3 both disappear. Instead, P6 is presence, indicating a nearby frequency that mode u is taken over by mode v. Therefore, the sequence of dominant mode is [R]-[u]-[v], indicating the situation that the zeros migrate into the complex plane as CMP-CNMP zeros. In other words, α_{01} is visualized by the sequence change from [R]-[u]-[R]-[v] to [R]-[u]-[v].

As α_0 further increases, one can expect that the sequence remains as [R]-[u]-[v]. However, when $\alpha_0>\alpha_{02}$, the zeros change from CMP-CNMP to two pairs of RMP-RNMP (dual RMP-

RNMP). Therefore, α_{02} cannot be visualized by the change of sequence. In other words, the sequence [R]-[u]-[v] only indicates that there are two zeros on the right half plane.

Next, a numerical example is presented. This example has the form as Eq. (3-9). We set α_u=-α_0, α_v=α_0, and use α_0=0.01, α_0=5, and α_0=21 to illustrate dual MMP, CMP-CNMP, and dual RMP-RNMP zeros respectively. Fig 3-12a demonstrates the Bode plot of the three cases, while Fig 3-12b shows their corresponding step responses. In Fig 3-12a, one can see that the difference between the case with [R]-[u]-[R]-[v] and the cases with [R]-[u]-[v] is that the latter ones both have significant large amount of contributions from mode u and mode v.

$$G(s)=\frac{1}{s^2+0.001s+1}+\frac{\alpha_u}{s^2+0.02s+10^2}+\frac{\alpha_v}{s^2+0.022s+11^2} \tag{3-9}$$

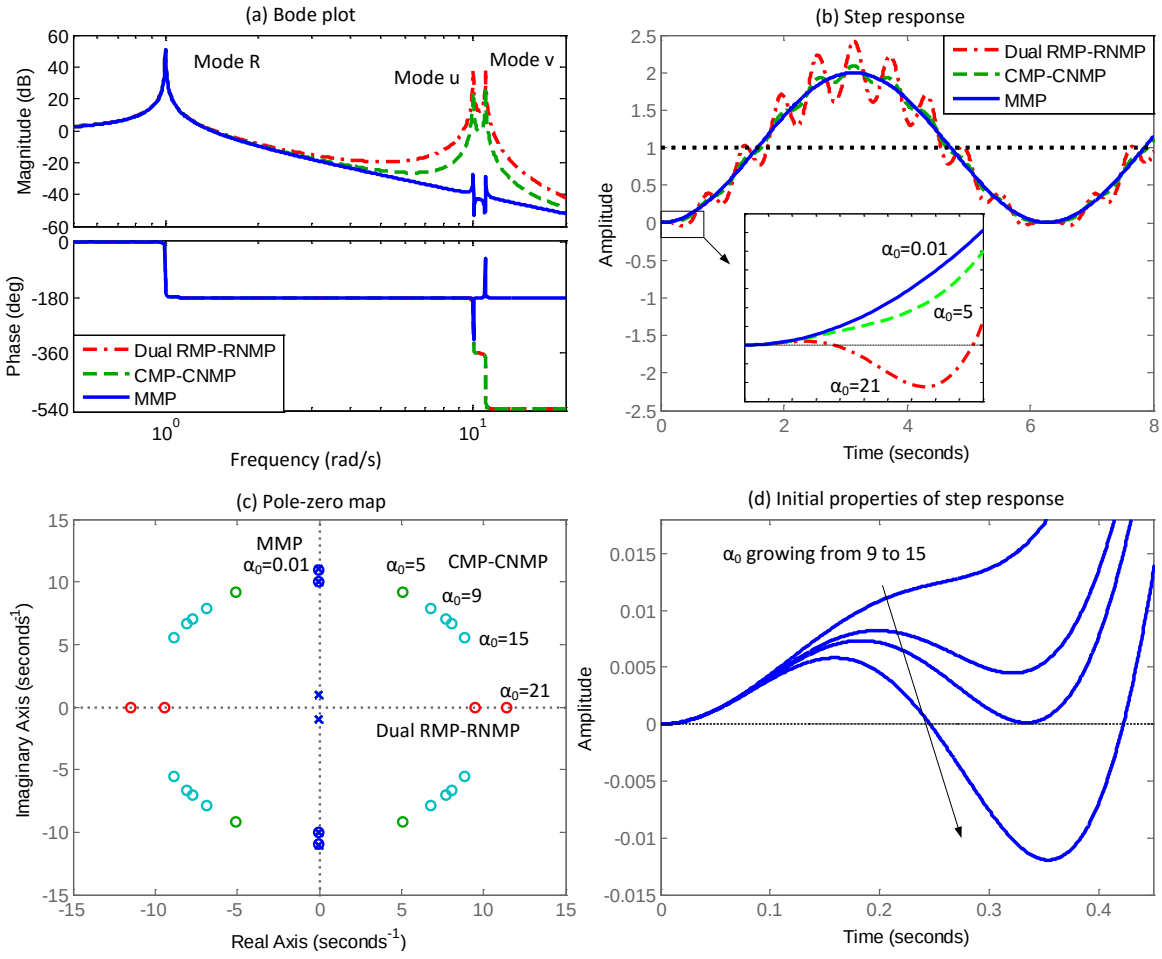

Fig 3-12 Numerical example of a system with CNMP zeros, under different α_0: (a) Bode plot; (b) step response; (c) pole-zero map; (d) initial part of step response.

For the dual MMP case, since $\alpha_R>0$ and $\alpha_v>0$, at high frequencies (i.e., $\omega>\omega_v$) mode v behaves similarly as mode R. Therefore, mode R is dominant mode in most of the frequency regions. As a result, the step response (Fig 3-12b blue solid) of the dual MMP case is essentially the step response of mode R (i.e., a second order undamped system in the positive direction).

For the dual RMP-RNMP case, the initial properties are discussed. Fast time-domain behavior is related to frequency-domain behavior at high frequencies. This is evident in the dual RMP-RNMP case (Fig 3-12b red dash-dot). We denote *f(t)* as the step response of *G(s*). Thus, the initial behavior is explained by examining the frequency-domain properties from high frequency to low frequency, which is the inverse direction of the sequence [R]-[u]-[v]. The first one in the inverse sequence is mode v. Consequently, *f(t)* starts with the step response of mode v, which is a second order undamped system in the positive direction. Second, mode v is inversely "taken over" by mode u. Thus *f(t)* exhibits a tendency of negative growing due to the step response of mode u, which is in the negative direction. The last one is mode R. As a result, *f(t)* eventually follows the step response of mode R. In addition, one can see that the number of direction reversal and the number of zero crossings are the same as the numbers of the RMP-RNMP pair. This observation is consistent with the previous literature [74].

For the CMP-CNMP case however, the variation of α_0 leads to different initial properties. Fig 3-12d demonstrates the step response of *G(s)* under different values of α_0. The corresponding migration of zeros is shown in Fig 3-12c. Note that all of these situations belong to the CMP-CNMP case. With different values of α_0, the step response exhibits a different number of zero crossings (i.e., zero, one, or two zero crossings), and different monotonousness (i.e., monotonous or non-monotonous). This observation is related to the study by Hoagg and Bernstein [73], who raised an open question on the relationship between the number of NMP zeros and the initial behavior in the step response of a system that contains these zeros. From Fig 3-12d, the number of NMP zeros in each case is the same while the initial behaviors are different, varying from the properties the same as MMP (dual) to those the same as RMP-RNMP (dual). Therefore, CMP-CNMP can be understood as a transition between these two types of zeros.

To explain the phase change, a comparison between a dual MMP case and the CMP-CNMP (or dual RMP-RNMP) case is illustrated in the Nyquist plot in Fig 3-13. Again, we use a simple settings as $\alpha_R=1$, $\alpha_v=\alpha_0>0$, and $\alpha_u=-\alpha_0$. One can see that the main characteristic for both cases is

the existence of two resonance circles (dual circles). For the dual MMP case, mode R (double integrator) "takes back over" at the medium frequency region (ω_u, ω_v). As a result, to the "corner" of the dual circles appears on the left side of the origin, leading to the phase rising twice. For the CMP-CNMP (or dual RMP-RNMP) case however, mode R does not "take back over" in (ω_u, ω_v). Consequently, the corresponding "corner" appears on the right side of the origin. Thus the phase drops twice (360°) due to the resonance circle without ever rising again. Also, one can see that $\alpha_R>0$, $\alpha_u<0$ and $\alpha_v>0$ play an important role for the 360° phase drop to happen.

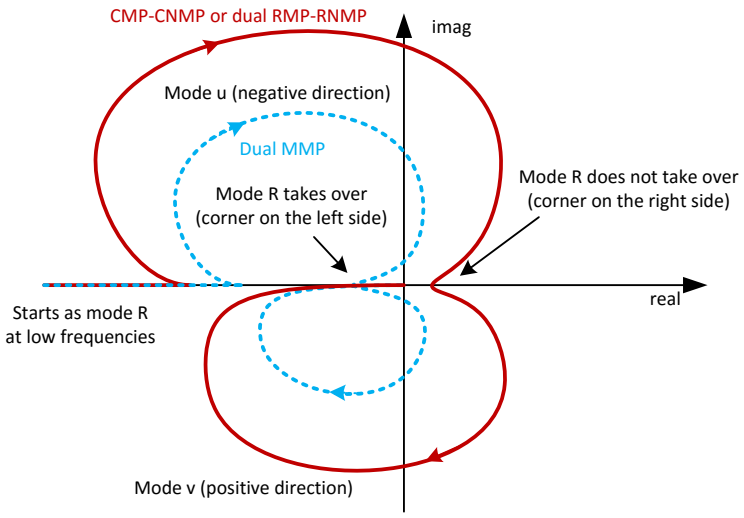

Fig 3-13 Nyquist plot of $G(s)$ showing the phase change ($\alpha_R>0$, $\alpha_u<0$ and $\alpha_v>0$).

Therefore, for the systems that described in Eq. (3-7) (under Assumptions A3-1, A3-2, A3-3, A3-6 and A3-7), CMP-CNMP (or dual RMP-RNMP) zeros appear if and only if the following two conditions are met. First, the three modes have alternating sign sequence (i.e., for any two adjacent modes, they are out-of-phase). Second, the sequence of dominant modes is [R]-[u]-[v] (i.e., each mode dominates at its corresponding frequency region).

The dual circles are more problematic than the single circle discussed in the RMP-RNMP case. Fig 3-14 illustrates the Nyquist plot of the open loop system that contains such dual circles. First, the dual circles exist in both the MMP case and the CMP-CNMP cases. Second, since the dual circles occupy all of the quadrants of the complex plane, it is almost impossible to "rotate" the dual circles to an angular position without endangering the critical point (-1, 0) (i.e., there is a

guaranteed peak in the sensitivity function). Moreover, if the damping ratio is reduced, the sizes of the dual circles increase dramatically, making the stability problem even more difficult. Similarly as Section 3.2.2, for a MMP case, it able to change the directions of the dual circles to the same direction by accepting the consequence of adding additional stable conjugate pole-zero pairs (an additional circle as shown in Fig 3-14). However, there may be some potential problems due to the small damping of the plant. This will be discussed in Chapter 6.

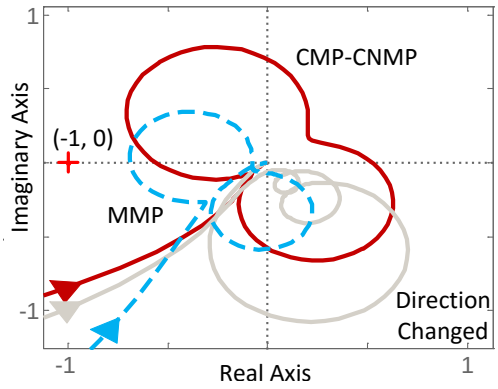

Fig 3-14 Nyquist plot of the open loop transfer functions (plant + controller) under different situations.

Table 3-3. The sequences of dominant modes and related zero types (double integrator + two modes).

Signs-Sequence	Sequence of dominant modes	Zero types	MMP Location
$\alpha_R>0$, $\alpha_u<0$, $\alpha_v>0$ (alternating)	[R]-[u]-[v]	CMP-CNMP or dual RMP-RNMP	N/A
	[R]-[u]-[R]-[v]	dual MMP	(ω_u, ω_v)
	[R]-[u]-[v]-[u]	MMP + RMP-RNMP	$(\omega_v, +\infty)$
	[R]-[v]-[u]-[v]	dual MMP	$(0, \omega_u)$
	[R]-[u]-[v]-[u]-[DI]	dual MMP	$(\omega_v, +\infty)$
$\alpha_R>0$, $\alpha_u>0$, $\alpha_v<0$	[Ru]-[v]	MMP + RMP-RNMP	$(0, \omega_u)$
	[Ru]-[v]-[DI]	dual MMP	$(0, \omega_u)$ and $(\omega_v, +\infty)$
$\alpha_R>0$, $\alpha_u<0$, $\alpha_v<0$	[R]-[uv]-[R]	dual MMP	(ω_u, ω_v) and $(\omega_v, +\infty)$
	[R]-[uv]	MMP + RMP-RNMP	(ω_u, ω_v)
$\alpha_R>0$, $\alpha_u>0$, $\alpha_v>0$	[Ruv]	dual MMP	$(0, \omega_u)$ and (ω_u, ω_v) (pole-zero alternating)

For a general situation not limited to the specific condition discussed above (i.e., α_R=1, α_v=α_0>0, and α_v=-α_0), a complete summary of the sequences is provided in Table 3-3. Note that the systems are limited to the types described as Eq. (3-7) (i.e., two modes with a double integrator). Similar as previous sections, we use [xy] to denote any sequence formed by mode x and mode y. Also, we use [DI] to denote an abstract dominant mode that behaves as a double integrator with positive a modal residue (i.e., α_D/s^2, and α_D>0). This is because in that case, the dominant mode is a combined effort of multiple modes and one cannot tell which one is the major contributor.

From Table 3-3, one can see that if CMP-CNMP and dual RMP-RNMP are regarded as a single class, the sequence of signs and the sequence of dominant modes can be used to uniquely map to the various situations of zeros, and this mapping is bijective.

Also, for cases other than CMP-CNMP, there are similarities between these cases and the results in Table 3-2 (or Table 3-1). In other words, these cases can be understood as applying the results in Table 3-2 twice. For example, consider the case when α_R>0, α_u>0, α_v<0 and [Ru]-[v]. For mode R and mode u (α_R>0, α_u>0), there must be a pair of MMP zeros, regardless of the sequence according to Table 3-2. Since mode v (α_v<0) only takes over at high frequencies, there must be a pair of RMP-RNMP zeros according to Table 3-2.

Therefore, the similarity between Table 3-3 and Table 3-2 implies that CMP-CNMP zeros refer to the case when a system as Eq. (3-7) cannot be decomposed into two simple systems listed in Table 3-2. The reason is that the two simple decomposed systems are not well-separated. This is consistent with the previous discussion that CMP-CNMP zeros can be understood as the transitions between dual MMP and dual RMP-RNMP.

In chapter 1, one of questions raised is: as CNMP zeros appear in the form of complex conjugate pair, how to evaluate the impact of CNMP zeros? Note that CNMP zeros appear in the form of CMP-CNMP zeros under Assumption A3-1 and A3-2. Since CMP-CNMP zeros are the transitions between dual MMP zeros and dual RMP-RNMP zeros, one can create a bound for the impact of CMP-CNMP zeros. In other words, CMP-CNMP zeros have weaker impact than dual RMP-RNMP zeros. This conclusion is consistent with the discussion in Fig 1-4.

3.2.5. Systems with Two Modes and a Constant

Following the last section, in this section we consider mode R as a mode much faster than mode u and mode v. As a result, Eq. (3-6) is restated as Eq. (3-10), where R is constant (i.e., the DC gain of mode R due to truncation).

$$G(s) = \frac{\alpha_u}{s^2 + \omega_u^2} + \frac{\alpha_v}{s^2 + \omega_v^2} + R;\ \omega_v > \omega_u > 0;\ \alpha_u > 0 \tag{3-10}$$

The analysis on the zeros of $G(s)$ in Eq. (3-10) is very similar as the last section. Thus we simply present the results as Table 3-4. Similar as the previous section, we use [DC] / [-DC] to denote an abstract dominant mode as the combinations of the DC gain of each mode in the positive / negative direction, respectively.

Table 3-4. The sequences of dominant modes and related zero types (two 2nd order modes + constant).

Signs-Sequence	Sequence of dominant modes	Zero types	MMP Location
$\alpha_u>0$, $\alpha_v<0$, $R>0$ (alternating)	[u]-[v]-[R]	CMP-CNMP or dual RMP-RNMP	N/A
	[u]-[R]-[v]-[R]	dual MMP	(ω_u, ω_v)
	[v]-[u]-[v]-[R]	MMP + RMP-RNMP	(0, ω_u)
	[DC]-[v]-[u]-[v]-[R]	dual MMP	(0, ω_u)
	[u]-[v]-[u]-[R]	dual MMP	(ω_v, $+\infty$)
$\alpha_u>0$, $\alpha_v<0$, $R<0$	[u]-[v]-[R]	MMP + RMP-RNMP	(ω_v, $+\infty$)
	[-DC]-[u]-[v]-[R]	dual MMP	(0, ω_u) and (ω_v, $+\infty$)
$\alpha_u>0$, $\alpha_v>0$, $R<0$	[R]-[uv]-[R]	dual MMP	(0, ω_u) and (ω_u, ω_v)
	[uv]-[R]	MMP + RMP-RNMP	(ω_u, ω_v)
$\alpha_u>0$, $\alpha_v>0$, $R>0$	[uvR]	dual MMP	(ω_u, ω_v) and (ω_v, $+\infty$) (pole-zero alternating)

3.3. Generalization Considerations

3.3.1. The Impact of Damping: a Simple Case

In this section, the impact of damping is examined. We still follow Assumptions A3-1 and A3-3. However, for Assumption A3-2, we assume small but non-zero damping. Consequently, the transfer function of a system can be written as Eq. (3-11).

$$G(s)=\sum_{i=1}^{n}\frac{\beta_i s+\alpha_i}{s^2+2\zeta_i\omega_i s+\omega_i^2} \tag{3-11}$$

Two main differences are observed between Eq. (3-11) and Eq. (3-2). First, each of the second order decomposed modes has its corresponding non-zero damping ratio (ζ_i). Second, each decomposed mode has an extra term ($\beta_i s$) in its numerator (i.e., modal residues). This extra term ($\beta_i s$) can be understood as the speed (i.e., first order derivative w.r.t. time) of the time-domain response of the i^{th} mode, with the amount of β_i.

In this section, we will discuss the impact of damping under a simple case. We assume $\beta_i=0$ for all of the decomposed modes. This is valid when the damping ratio is sufficiently small. Next, we start with the variation of modal residues of only two modes. Similar to the discussion on RMP-RNMP zeros in Section 3.2.3, Eq. (3-5) can be extended to Eq. (3-12).

$$G(s)=\frac{\alpha_u}{s^2+2\zeta_u\omega_u s+\omega_u^2}+\frac{\alpha_v}{s^2+2\zeta_v\omega_v s+\omega_v^2};\ \omega_u<\omega_v \tag{3-12}$$

From Eq. (3-12) one obtains Eq. (3-13).

$$G(s)=\frac{\alpha_u/\alpha_v(s^2+2\zeta_v\omega_v s+\omega_v^2)+(s^2+2\zeta_u\omega_u s+\omega_u^2)}{(s^2+2\zeta_u\omega_u s+\omega_u^2)(s^2+2\zeta_v\omega_v s+\omega_v^2)/\alpha_v}\triangleq\frac{(\alpha_u/\alpha_v)\cdot A(s)+B(s)}{D(s)} \tag{3-13}$$

Similar as previous sections, the locus technique can be applied. Fig 3-15 demonstrates the zero loci of *G(s)* with a non-zero damping ratio. For ease of representation, we only show the situation when $\alpha_u\cdot\alpha_v<0$ (i.e., out-of-phase).

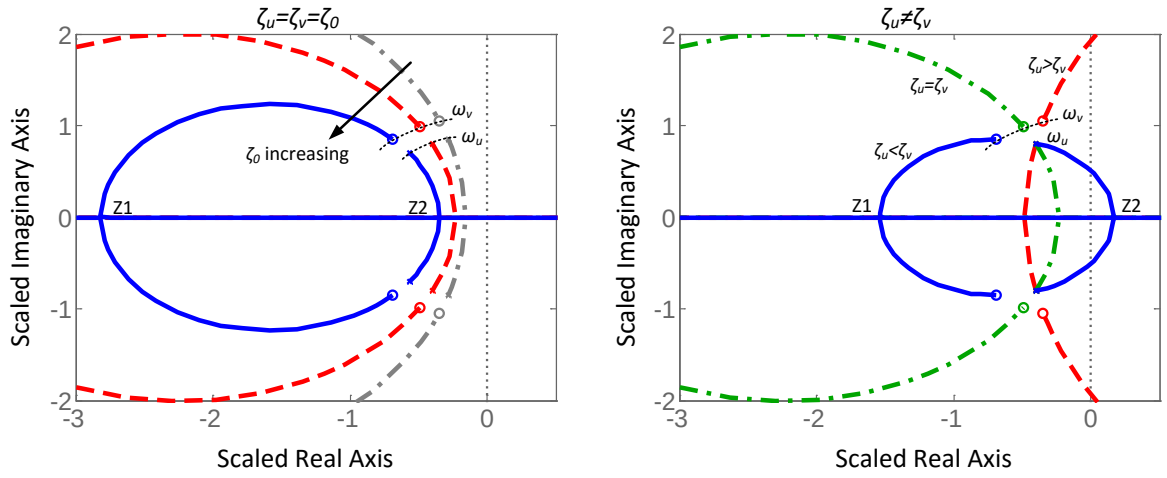

Fig 3-15 Migration of zeros with non-zero damping ratios (only two modes): (a) damping ratios are the same; (b) damping ratios are different.

Three main observations are listed, in contrast to Fig 3-8. First, the poles of mode u and mode v are shifted to the left by a distance that is determined by the amount of damping. Second, the zeros are still presented in the form of a pair. However, these zeros are no longer anchored on the imaginary axis. Additionally, when these zeros appear on the real axis, they are no longer symmetric w.r.t. the imaginary axis (though they are symmetric w.r.t. the zero loci in some sense). More importantly, when $\zeta_u > \zeta_v$, the circle-like locus is repelled to right hand side. In this case, the pair of zeros may become CNMP zeros. In other words, if the slower mode u has a larger damping ratio than the faster mode v, then it is possible that a pair of CNMP zeros appear. Third, both the break in/away points (Z1 and Z2) are finite, with their positions governed by Eq. (3-14).

$$\left(\zeta_u\omega_u - \zeta_v\omega_v\right)s^2 + \left(\omega_u^2 - \omega_v^2\right)s + \omega_u\omega_v\left(\zeta_v\omega_u - \zeta_u\omega_v\right) = 0 \tag{3-14}$$

Regarding the CMP-CNMP case, we present a simplified situation as an extension of Eq. (3-7) as shown in Eq. (3-15). For simplicity, mode u and mode v are assumed to have the same amount of damping ratio (ζ_0).

$$G(s)=\frac{1}{s^2}+\frac{\alpha_u}{s^2+2\zeta_0\omega_u s+\omega_u^2}+\frac{\alpha_v}{s^2+2\zeta_0\omega_v s+\omega_v^2};\ \omega_u<\omega_v \tag{3-15}$$

Fig 3-16 demonstrates a numerical example (ω_u=0.9, ω_v=1.1, and α_u/α_v=2) with different values of damping ratios. It is noteworthy that the loci change with varying the ratio of modal residues (α_u/α_v). Regarding the case when $\alpha_u/\alpha_v > \omega_v^2/\omega_u^2$, it can be seen that all of the loci move to the left as the damping ratio increases. Moreover, there is a certain value of the damping ratio when all of the loci are located on the left side of the complex plane. This damping ratio can be obtained by evaluating the crossing of the loci w.r.t. the imaginary axis, as shown in Eq. (3-16).

$$\zeta=\frac{1}{2}\sqrt{\frac{\kappa-1}{\kappa\eta^{-1}-\eta}+\frac{\kappa\eta^{-1}-\eta}{\kappa-1}-\eta-\eta^{-1}};\quad \eta=\frac{\omega_u}{\omega_v}<1;\quad \kappa=\frac{\alpha_u}{\alpha_v}>1 \tag{3-16}$$

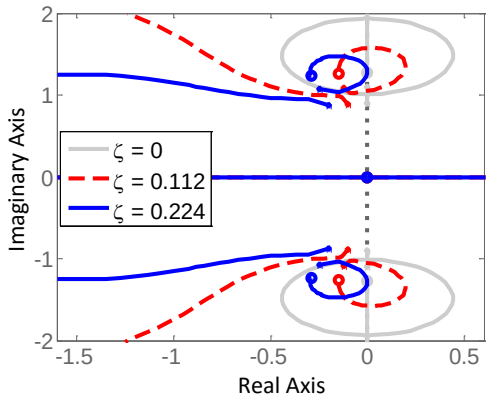

Fig 3-16 Migration of zeros under different damping ratios (two modes with a remainder, $\alpha_u/\alpha_v > \omega_v^2/\omega_u^2$).

It is important to note that since the damping is non-zero, the phase of a system is no longer fixed to either positive (i.e., 0°) or negative (i.e., 180°). Instead, there will be transitions between the two directions. As a result, the definition of mode taking over, the dominant mode, and thus the sequence of dominant mode may not apply to the situation with large damping. The impact of damping for the technique of sequence is a future work worth investigating.

3.3.2. Other Considerations

In addition to the representative but specific situations discussed in the above sections, the presented work can be extended to cover more general cases. The extensions are based on removing the assumptions made previously, and some are briefly discussed as below.

First, the work in this chapter is based on simple systems (e.g., systems with only two modes). For systems with multiple modes, how to apply the developed results is of great interest. Second, one can extend the cases of two closely spaced modes to multiple closely spaced modes. This extended analysis represents a broader class of module-based systems. Third, continuous system can be considered. Lastly, an even broader range of systems can be included if the decomposed modes are first order systems or a hybrid form of first order and second order systems. In these situations, the zeros may no longer appear in the form of pairs.

3.4. Contributions and Conclusions

The key contributions of this chapter are concluded as below. First, based on the variation of modal residues, a systematic step by step investigation is provided to explain the zeros of lightly damped flexible systems. Under Assumption A3-1 and A3-2, the zeros appear in the form of pairs. Particularly for NMP zeros, RNMP zeros appear in the form of RMP-RNMP zeros and CMP zeros appear in the form of CMP-CNMP zeros. Furthermore, the minimum system that has CNMP zeros is a system with two closely spaced modes and a remainder. Second, a tool is proposed using the sequence of dominant modes and the sequence of their signs to interpret the variation of zero types. For a system with only two modes, RMP-RNMP zeros appear if and only if the two modes are out-of-phase and the faster mode dominates the high frequency region. For a system with two modes and a remainder, we demonstrate a similarity between CMP-CNMP zeros and dual RMP-RNMP zeros. Furthermore, these two types of zeros appear if the decomposed modes have alternating signs and dominate under a certain sequence. Moreover, this tool allows us to understand the initial properties of the step response of systems with NMP zeros. In addition, the difficulties that RMP-RNMP zeros and CMP-CNMP zeros bring in for the close

loop control are visualized. One can relieve these difficulties albeit with certain tradeoffs. Third, the investigation is extendable to a more general form, such as non-zero damping.

In addition, there are several new open questions posed in this work that are being addressed by us. First, it is desirable to generalize the framework. Second, the variation parameters in this framework are essentially mathematical parameters. Greater insights into the relationship between these parameters and physical parameters are desirable. Last, based on the findings of this chapter, physical system and control system design can be explored to improve the dynamic performance of flexible systems over the state-of-the-art level.

Chapter 4. On the Correlation between CMP-CNMP Zeros and Curve Veering

In the work of the previous chapters, Chapter 2 presents an approach based on physical parameters to model when the CNMP zeros appear (in the form of CMP-CNMP quartet) observed in experiments. Chapter 3 presents a mathematical model that utilizes the sequence of dominant modes to reveal a sufficiency condition for the existence of CMP-CNMP zeros and an associated physical interpretation. The variation parameter in the former study is the operating point while the variation parameters in the latter study are the modal residues. However, the mapping between the operating point and the modal residues has not been shown in the previous chapters, resulting in a correlation gap between the physical model and the mathematic model.

To address this gap, the phenomenon of curve veering is introduced in this chapter. It is shown that the existence of curve veering under certain conditions will guarantee the specific sequence of dominant modes and their signs that leads to CMP-CNMP zeros. Therefore, this finding validates the hypothesis proposed in Section 2.3 (i.e., there is a correlation between CNMP and curve veering). Furthermore, by combining the results from Chapters 2 to 4, we demonstrate the transition from MMP to CMP-CNMP zeros for certain flexure mechanisms (i.e. how CNMP zeros appear).

The motivation of introducing the phenomenon of curve veering into this work is due to the fact that CNMP zeros and curve veering share the same key attributes. From the discussion in Chapter 2, three key factors are shown to be closely related to the existence of CMP-CNMP zeros in flexure mechanisms. First, there are multiple modes with their natural frequencies closely spaced. This is due to the fact that the flexure mechanisms usually employ multiple sub-structures that are based on the same module or building block. Moreover, all of these sub-structures have nominally the same dimensions (i.e., these sub-structures are periodic/symmetric). Second, there can be certain parametric asymmetry between these periodic/symmetric sub-structures that may arise either intentionally (e.g., physical parameter choices) or unintentionally

(e.g., manufacturing tolerances). Third, there is certain mechanical structure that connects these sub-structures (e.g., Fig 2-2).

Interestingly, the phenomenon of curve veering shares similar key attributes as those mentioned above. One of the earliest studies in curve veering was conducted by Leissa [53] in 1974, while early studies of CNMP zeros include Tohyama and Lyon [48] in 1989 as well as Loix et al. [49] in 1996. If these two fields that have been studied separately for decades are shown to be connected on certain conditions, then we can extend our understanding on both phenomena one step further. Moreover, since curve veering widely exists in periodic/symmetric structures, this work can be extended to a broader range of flexible systems not just limited by module-based (e.g., DPFM) flexure mechanisms.

4.1. Literature Review

Curve veering refers to the situation when the eigenvalue loci of two closely spaced modes in a system approach each other and then diverge, as a result of certain parametric variation [53]. The point in the parametric space where these two modes have the closest natural frequencies is called the veering point. Specifically, the situation that occurs when these two modes intersect (i.e., when they have the same natural frequency) at the veering point is called mode crossing [54]. In some references, curve veering is termed as mode veering or mode coalescence [82].

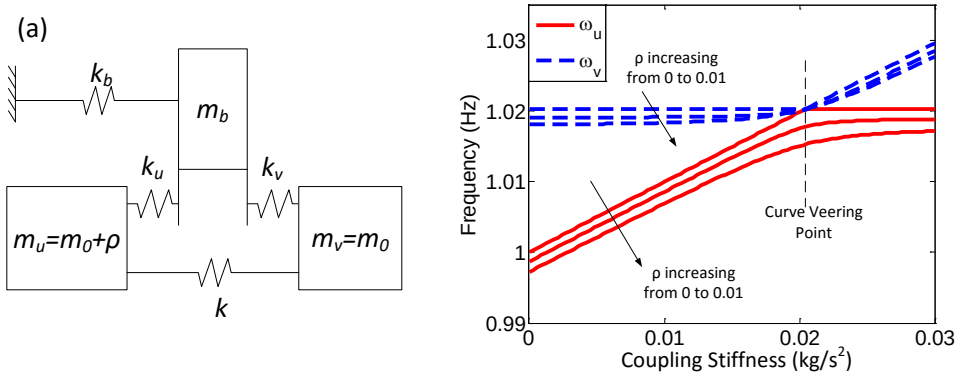

Fig 4-1 3 DoFs spring mass model example: (a) model illustration; (b) the loci of natural frequencies.

A simple three degree-of-freedom (3 DoF) example is demonstrated in Fig 4-1a. This model is enlightened by the traditional 2 DoF spring-mass models in the studies of curve veering [83]. To illustrate the eigenvalue loci, we use the numerical setting as follows: $k_1=k_2=k_b=1$ kg/s^2, $m_b=50$ kg, $m_u=1$ kg, and $m_v=1+\rho$ kg. For the variation of parameters, we set k varying from 0 to 0.03 kg/s^2, and ρ varying from 0 to 0.01 kg. The loci of the natural frequencies associated with the modes dominated by m_u and m_v are shown in Fig 4-1b and the above-defined curve veering is exhibited. Moreover, mode crossing (i.e., the situation when $\omega_u=\omega_v$) is observed when $\rho=0$ (i.e., when the two masses m_u and m_v are symmetric).

Curve veering can be mathematically understood as a situation when the system is very close to the status of having repeated eigenvalues [55, 84, 85]. In the vicinity of the veering point, eigenvectors undergo dramatic changes [86]. These dramatic changes are recognized as mode localization [56, 87, 88]. In addition, the system could become so-called “critically configured” indicating a sensitive situation in which small changes in a system parameter could cause large changes in system response [55].

The phenomenon of curve veering is studied in situations where there are symmetric or periodic structures, such as blade disks and turbine blades [89-92], wind-bridges [93], and module-based flexure mechanisms [72]. These situations can be concluded as flexible systems with nominally the same dimensional modules, which lead to multiple closely spaced modes. The studies of curve veering can be applied to the safety evaluation [91]. There are also techniques using curve veering to achieve vibration isolation [94].

There can be various selections of the varying parameter, such as the mass value [83], the stiffness value [87], and other operating conditions [72]. The large number of different choices makes each study case-specific and thus limits the referential value. However, the conclusions of these studies are inherently consistent. The reason is that all of the different choices fall into the same two of the key attributes of curve veering: (1) the connection (coupling) between the nominally symmetric/periodic structures, and (2) the parametric asymmetry between these structures. Therefore, in this chapter the varying parameter is selected based on the principle that it mainly affects one of the above two attributes without significantly affecting the other. For example, in the model shown in Fig 4-1a, the coupling spring stiffness k mainly affects the

spring connecting between the two nominally symmetric masses (m_u and m_v), while the mass asymmetry ρ mainly affects the balance between these two masses.

Based on the above discussion, one can conclude that there are totally three key attributes in curve veering. First, there are multiple nominally periodic/symmetric sub-structures within the given system. Second, there are certain asymmetries between these sub-structures. Third, there is certain coupling between these sub-structures. It is noteworthy that these attributes are the necessary conditions of curve veering. However, the sufficiency of these attributes is rarely discussed in the literature to the best of our knowledge. Moreover, one can see the strong similarity between the key attributes of CNMP zeros and those of curve veering, which motivates the study in this chapter.

4.2. The Correlation between CMP-CNMP Zeros and Curve Veering

4.2.1. Preliminary Assumptions

This section has certain overlap with Section 3.2. But we recall all of the necessary equations and assumptions for ease of discussion in this chapter. Regarding the scope of discussion in this chapter, the systems considered are linear time invariant (LTI) and single-input-single-output (SISO). Assumptions made in this chapter are listed as below.

Assumption A4-1: the system is a structural system. In other words, all of the decomposed modes are second order and there are no first order modes.

Assumption A4-2: damping is negligible.

Assumption A4-3: the system has no right half plane poles (i.e., the plant is open loop stable).

Assumption A4-4: consider the transfer function *G(s)* of a system and the two modes of interest (denoted as mode u and mode v). For frequencies near the range (ω_u, ω_v), the contributions of modes other than mode u and mode v are dominated by a single mode (denoted as mode R). Moreover, mode R is well-separated from mode u and mode v.

Assumption A4-5: consider mode R as a double integrator.

Consequently, a system under Assumption A4-1 to A4-5 is stated as below. Without losing generality, mode u is defined to be slower than mode v.

$$G(s)=\frac{\alpha_R}{s^2}+\frac{\alpha_u}{s^2+\omega_u^2}+\frac{\alpha_v}{s^2+\omega_v^2};\ \omega_v>\omega_u>0 \tag{4-1}$$

Next, *G(s)* is written into the rational polynomial form as

$$G(s)=\frac{N(s)}{D(s)}=\frac{\left[\alpha_v+\alpha_u+\alpha_R\right]s^4+\left[\left(\omega_u^2+\omega_v^2\right)\alpha_R+\alpha_u\omega_v^2+\alpha_v\omega_u^2\right]s^2+\alpha_R\omega_u^2\omega_v^2}{s^2\left(s^2+\omega_u^2\right)\left(s^2+\omega_v^2\right)} \tag{4-2}$$

Note that in this section there is no requirement of the curve veering phenomenon. In the next two sections, we will first examine the necessary and sufficient conditions of CMP-CNMP, and then bring in the correlation between CMP-CNMP and the curve veering phenomenon.

4.2.2. The Parabola of CMP-CNMP Zeros

Based on Eq. (4-2), if we define $x=s^2$, the numerator of *G(s)* is shown below.

$$\begin{aligned}N(x)&=\left[\alpha_v+\alpha_u+\alpha_R\right]x^2+\left[\left(\omega_u^2+\omega_v^2\right)\alpha_R+\alpha_u\omega_v^2+\alpha_v\omega_u^2\right]x+\alpha_R\omega_u^2\omega_v^2\\&=b_2x^2+b_1x+b_0\end{aligned} \tag{4-3}$$

From Eq. (4-3) one can see that *N(x)* is a second order polynomial as a function of *x*, which is a parabola. The condition *N(x)*=0 corresponds to any crossing(s) of the parabola with the horizontal line. This gives the roots of *N(x)* in terms of *x* and thus the zeros of *G(s)* in terms of *s*. If there is no crossing between the parabola and the horizontal line, then the roots of *N(x)* are complex conjugate numbers. As a result, *G(s)* has CMP-CNMP quartet zeros.

To illustrate, a specific numerical example is provided. If we vary α_R, while fixing other variables as $\alpha_u=-1$, $\alpha_v=1$, $\omega_u^2=1$, and $\omega_v^2=2$, then the second order coefficient b_2 in Eq. (4-3) is solely affected by α_R, which determines the "flatness" of the parabola. Fig 4-2a demonstrates the different situations of crossing between the parabola and the horizontal line under different values of α_R.

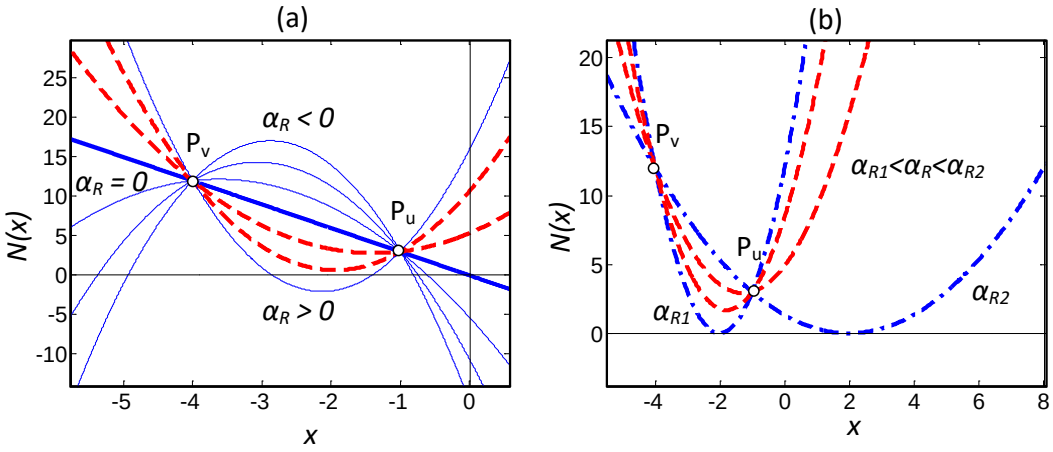

Fig 4-2 Numerical example of parabola $N(x)$ under different values of α_R (α_u=-1, α_v=1, ω_u^2=1, and ω_v^2=2). Dashed lines indicate the situations with CNMP zeros. (a) α_R varies from negative to positive; (b) α_R is negative and varies between two critical values.

In this numerical example, situations with different signs of α_R in Fig 4-2a are discussed as below. First, when α_R<0, the opening of the parabola faces downward. Therefore, there must be two crossings between the parabola and the horizontal line. Moreover, since the crossings happen at x<0, then by $s^2=x$, this situation corresponds to a pair of marginally minimum phase zeros on the imaginary axis, and they are symmetric w.r.t. the real axis. Second, when α_R=0, the parabola reduces to a line passing through the origin. This case refers to the situation where the mode R is negligible. Third, when α_R>0, the opening of the parabola faces upward. Moreover, as the absolute value of α_R increases, the parabola becomes less "flattened". Therefore, there exist instances when the parabola has no crossing with the horizontal line, which is illustrated as the dashed lines in Fig 4-2a. When there is no crossing, solving $N(x)$=0 results in complex roots of x, indicating that the zeros of $G(s)$ are in the form of CMP-CNMP zeros (quartet).

It is also noteworthy that there are two points (denoted as P_u and P_v in Fig 4-2a) shown to be invariant of α_R. This can be verified by substituting $x=-\omega_u^2$ and $x=-\omega_v^2$ into $N(x)$ in Eq. (4-3) as shown below. Note that there are no specific numerical values required.

$$P_u:\left[-\omega_u^2,-\alpha_u\omega_u^2\left(\omega_v^2-\omega_u^2\right)\right];\ P_v:\left[-\omega_v^2,\alpha_v\omega_v^2\left(\omega_v^2-\omega_u^2\right)\right] \tag{4-4}$$

The existence of these two points is explained as below, using P_u as an example. When x is closed to $-\omega_u^2$, the situation indicates that $G(s)$ is evaluated at frequencies closed to the natural frequency of mode u. As a result, the magnitude of $G(s)$ is dominated by mode u, while the mode R and mode v are negligible. Therefore, α_R and α_v are not presented in the vertical value of P_u in Eq. (4-4). Similarly, α_R and α_u are not presented in the vertical value of P_v. Thus, P_u and P_v are invariant of α_R.

From Eq. (4-4), since we define $\omega_v>\omega_u$, the signs of the vertical values of P_u and P_v are determined by the signs of α_u and α_v. Moreover, from Fig 4-2a one can see that for CMP-CNMP to appear, P_u and P_v have to be either above the horizontal line with $\alpha_R>0$ (i.e., $\alpha_R>0$, $\alpha_u<0$, and $\alpha_v>0$), or below the horizontal line with $\alpha_R<0$ (i.e., $\alpha_R<0$, $\alpha_u>0$, and $\alpha_v<0$). This indicates that α_R, α_u, and α_v form a sequence of alternating signs.

Next, if P_u and P_v are on the same side of the horizontal line (i.e., $\alpha_u\cdot\alpha_v<0$), one can calculate the corresponding values of the second order coefficient (b_2), referring to the instance when the parabola $N(x)$ has a single crossing with the horizontal line (i.e., condition of tangency; see the dashed-dot lines in Fig 4-2b). Accordingly, the critical values of α_R can be obtained as below. Note that for values between α_{R1} and α_{R2}, the parabola has no crossing with the horizontal line, which indicates the presence of CMP-CNMP zeros in $G(s)$.

$$\alpha_{R1,R2}=\left(\alpha_v\omega_u^2-\alpha_u\omega_v^2\pm 2\omega_u\omega_v\sqrt{-\alpha_u\alpha_v}\right)/\left(\omega_v^2-\omega_u^2\right);\ \alpha_u\cdot\alpha_v<0 \tag{4-5}$$

Accordingly, the corresponding values of the second order coefficient b_2 is obtained by the relationship $b_2=\alpha_R+\alpha_u+\alpha_v$ as below. Note that when $\alpha_u<0$ and $\alpha_v>0$, b_{21} and b_{22} are positive, and vice versa. For ease of presentation, we define $|b_{22}|>|b_{21}|$.

$$b_{21,22}=\operatorname{sign}\left(\alpha_v\right)\cdot\left(\omega_v\sqrt{|\alpha_v|}\pm\omega_u\sqrt{|\alpha_u|}\right)^2/\left(\omega_v^2-\omega_u^2\right);\ \alpha_u\cdot\alpha_v<0 \tag{4-6}$$

It is noteworthy that the results in Eq. (4-6) are related to the two break in/away points in the locus of zeros such as the example shown in Fig 3-9.

4.2.3. Normalization and the Propositions

In last section, the crossing between a parabola associated with the system and the horizontal line is used to examine the presence of CMP-CNMP zeros. Regarding this parabola, two points (P_u and P_v) that are invariant of α_R are shown to be useful in checking the crossing situations. When these two points are fixed, the condition causing CMP-CNMP zeros is described by a specific range of values of the second order coefficient (b_2). In this analysis, α_R is chosen as the varying parameter for simplicity. However, if mode u and mode v are under curve veering, the eigenvectors associated with other modes are assumed to be invariant (Assumption A4-4). As a result, the modal residue of mode R is unchanged in the vicinity of curve veering (i.e., α_R is constant). Instead, there are multiple parameters under variation (i.e., α_u, α_v, ω_u, and ω_v), which complicate the analysis of zeros. Therefore, a normalization (or scaling) method is proposed to transform the complex situation of curve veering to a simpler case.

The normalization method is shown in Eq. (4-7).

$$x = (\omega_v^2 - \omega_u^2)\hat{x}; \quad N = \omega_u^2(\omega_v^2 - \omega_u^2)\hat{N} \tag{4-7}$$

P_u and P_v are then scaled to $\hat{P}_u$ and $\hat{P}_v$, with their coordinates ($\hat{x}$, $\hat{N}$) shown as Eq. (4-8).

$$\hat{P}_u : \left[-\omega_u^2/\left(\omega_v^2 - \omega_u^2\right), -\alpha_u\right]; \ \hat{P}_v : \left[-\omega_v^2/\left(\omega_v^2 - \omega_u^2\right), \ \alpha_v\omega_v^2/\omega_u^2\right] \tag{4-8}$$

Based on Eq. (4-7), one can see that after normalization, the horizontal distance between the two points $\hat{P}_u$ and $\hat{P}_v$ is scaled to unity. Moreover, the second order coefficient is scaled as follows, where $\omega_\Delta = \omega_u - \omega_v$.

$$\hat{b}_2 = \left(\omega_v^2 - \omega_u^2\right)b_2/\omega_u^2 = \omega_\Delta\left(\omega_v + \omega_u\right)\left(\alpha_u + \alpha_v + \alpha_R\right)/\omega_u^2 \tag{4-9}$$

Additionally, the critical values for the second order coefficient (i.e., b_{21} and b_{22}) can be scaled as follows.

$$\hat{b}_{21,22} = \frac{1}{\omega_u^2}\cdot\text{sign}\left(\alpha_v\right)\cdot\left(\omega_v\sqrt{|\alpha_v|} \pm \omega_u\sqrt{|\alpha_u|}\right)^2; \ \alpha_u\cdot\alpha_v < 0 \tag{4-10}$$

Note that for $\hat{b}_{21}$ and $\hat{b}_{22}$ to exist, one requires $\alpha_u \cdot \alpha_v < 0$. Similarly, we define $|\hat{b}_{22}| > |\hat{b}_{21}|$. Next, consider the following statements.

Condition C4-1: α_R, α_u, and α_v form a sequence of alternating signs.

Condition C4-2: $\hat{b}_{21}$ and $\hat{b}_{22}$ exist, and either Eq. (4-11) or Eq. (4-12) is met.

$$\hat{b}_{22} > \hat{b}_2 > \hat{b}_{21} > 0;\ \alpha_u < 0;\ \alpha_v > 0 \tag{4-11}$$

$$\hat{b}_{22} < \hat{b}_2 < \hat{b}_{21} < 0;\ \alpha_u > 0;\ \alpha_v < 0 \tag{4-12}$$

Proposition P4-1: for systems under Assumptions A4-1 to A4-5, CMP-CNMP zeros appear if and only if Condition C4-1 and C4-2 are met (necessary and sufficient conditions).

The proof of Proposition P1 is shown in the above procedure of finding the two critical situations of the parabola family (see Fig 4-2a and Fig 4-2b). Note that by far the presence of curve veering is not required. Next, we bring in the scenario of curve veering. Since it is complicated to directly examine the correlation between the phenomenon of CMP-CNMP and the phenomenon of curve veering, we will examine how the presence of curve veering affects to the necessary and sufficient conditions of CMP-CNMP zeros.

Condition C4-3: the two closely spaced modes (i.e., mode u and mode v) experience curve veering with certain parameter (denoted as ψ) variation, and their natural frequencies are sufficiently close at the veering point ψ_{cv} (i.e., $\omega_\Delta \to 0$, where $\omega_\Delta = \omega_u - \omega_v$).

It is known that there have to be at least two modes for the existence of curve veering. Moreover, the number of modes related to curve veering is the number of periodic structures (e.g., bladed disks). In this chapter, we focus on the situation with only two modes that are related to curve veering, while situation with multiple modes can be extended analogically. The description “sufficiently close” in Condition C4-3 indicates that the parameter set of the system is almost the parameter set of mode crossing.

Condition C4-4: $\hat{b}_{21}$ and $\hat{b}_{22}$ exist, and there exists a value $\psi_0 \neq \psi_{cv}$ such that either Eq. (4-13) or Eq. (4-14) is met.

$$\hat{b}_2 > \hat{b}_{21} > 0;\ \alpha_u < 0;\ \alpha_v > 0 \tag{4-13}$$

$$\hat{b}_2 < \hat{b}_{21} < 0;\ \alpha_u > 0;\ \alpha_v < 0 \tag{4-14}$$

Proposition P4-2: for systems under Assumptions A4-1 to A4-5, if Condition C4-1, C4-3, and C4-4 are met, then there exists a value $\psi=\psi_x$ such that CMP-CNMP zeros appear.

To prove, if condition C4-3 is met, then one has $\omega_\Delta \to 0$ at the curve veering point. Since all of the modal residuals (i.e., α_u, α_v, and α_R) are finite, $\hat{b}_2$ is finite (see Eq. (4-9)), then $\hat{b}_2 \to 0$ as $\omega_\Delta \to 0$. Based on Condition C4-3, if Condition C4-1 and C4-4 are met, then in the variation of ψ, there exists an instance (denoted as ψ_x) when $\hat{b}_2$ is in the range of ($\hat{b}_{21}$, $\hat{b}_{22}$) due to continuity, which indicates Condition C4-2 is met. Therefore, based on Proposition P4-1, CMP-CNMP zeros are presence when $\psi=\psi_x$.

One can see that Condition C4-4 has a weaker requirement than Condition C4-2, because of the presence of curve veering. In other words, in the vicinity of curve veering (Condition C4-3), the requirement for CMP-CNMP zeros is weakened from Condition C4-2 to Condition C4-4.

Furthermore, it can be seen that Condition C4-1, C4-3, and C4-4 are the sufficient but not necessary conditions of the presence of CMP-CNMP zeros. In other words, Proposition P4-2 indicates that at the presence of CMP-CNMP zeros, the phenomenon of curve veering may not exist. An example is shown in Fig 4-3, which is inspired by the modal decomposition in the literature [59]. In Fig 4-3, all masses can only move in the vertical direction. The contribution of each spring-mass (i.e., subsystems with the subscript as R, u, and v) to the output *y* is determined by their associated lever arms. Moreover, each spring-mass is isolated from the other ones. Therefore, the output *y* can be any linear combination of these spring-mass modes. Now consider applying the same scalar force *F* on each of the masses and examine the transfer function from *F* to *y*. It can be seen that with a certain selection of parameters (i.e., masses, springs, and the length of the lever arms), there will be CMP-CNMP zeros in this transfer function. However, since there is no coupling between mode u and mode v, there is no curve veering in this system.

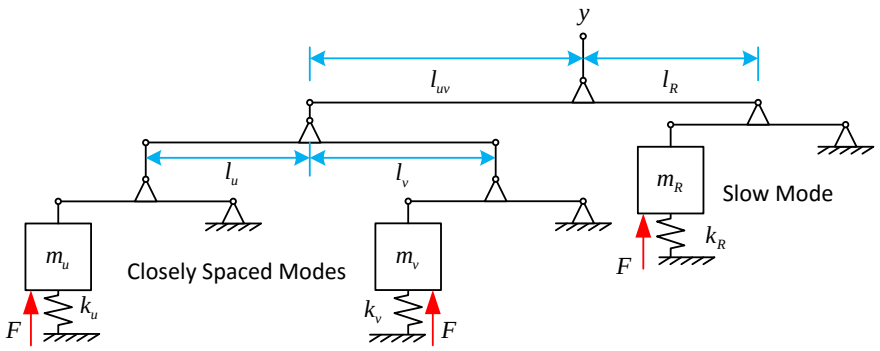

Fig 4-3 Spring-mass system with lever arms.

However, for systems under Assumptions A4-1 to A4-5, if Condition C4-3 is always true, then Condition C4-1 and C4-4 become the necessary and sufficient conditions of the presence of CMP-CNMP zeros.

Proposition P4-2x: for systems under Assumptions A4-1 to A4-5, if Condition C4-3 is true, then CMP-CNMP zeros appear if and only if Condition C4-1 and C4-4 are met.

Compared Proposition P4-2x with Proposition P4-1, since Condition C4-4 is a weaker requirement of Condition C4-2, one can see that the presence of curve veering makes CMP-CNMP zeros more likely to appear.

Next, since Condition C4-4 is related to Eq. (4-9) and Eq. (4-10), to further understand the physical insights of Condition C4-4, a simpler case of Condition C4-4 is examined next, which is a subset of Condition C4-4.

Condition C4-4y: in the variation of ψ, there exists an instance of ψ when there are two pairs of MMP zeros (termed as dual MMP zeros) in the range (ω_u, ω_v).

Note that Condition C4-4y is considered for systems under Assumptions A4-1 to A4-5. In addition, Condition C4-4y implies Condition C4-1 to be true. This can be shown using the parabola in Eq. (4-3). Condition C4-4y indicates that the two points invariant of α_R (i.e., P_u and P_v) are on the same side of the horizontal, which implies $\alpha_u \cdot \alpha_v < 0$. Moreover, the parabola has two crossings with the horizontal, and these two crossings are between the horizontal positions of P_u and P_v. Therefore, P_u and P_v have to be either above the horizontal line with $\alpha_R > 0$ (i.e.,

α_R>0, α_u<0, and α_v>0), or below the horizontal line with α_R<0 (i.e., α_R<0, α_u>0, and α_v<0). This indicates that α_R, α_u, and α_v form a sequence of alternating signs.

Furthermore, since the parabola has two crossings between the horizontal positions of P_u and P_v, based on Eq. (4-10), one has $\hat{b}_2 > \hat{b}_{22}$, which implies that Condition C4-4 is met. Therefore, Condition C4-4y is a subset of Condition C4-4. Next, we have a corollary of Proposition P4-2.

Proposition P4-2y: for systems under Assumptions A4-1 to A4-5, if Condition C4-3 and C4-4y are met, then there exists a value $\psi=\psi_x$ such that CMP-CNMP zeros appear (sufficient but not necessary conditions).

The situation described by Condition C4-4y can be further explained using the technique named "sequence of the dominant modes" in Chapter 3. For a system under Assumption A4-1 to A4-5, Condition C4-4y indicates a situation demonstrated in Fig 4-4a, where mode R is the double integrator, mode u and mode v are the two modes related to curve veering.

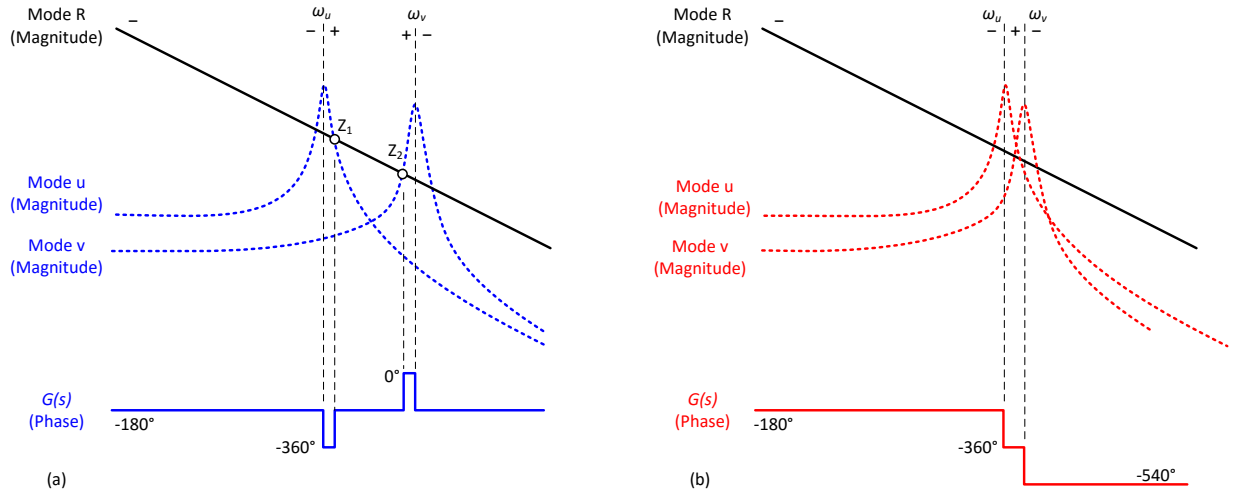

Fig 4-4 Change of sequence of the dominant mode due to curve veering. (a) dual MMP zeros; (b) CMP-CNMP (or dual RMP-RNMP) zeros

Without losing generality, we consider the alternating signs as α_R>0, α_u<0, and α_v>0. Therefore, the sign of mode R is always negative (i.e., −180° in phase). Since α_u<0, there is a crossing between mode R and mode u at Z_1, where mode R and mode u have the same magnitude but different signs. As a result, Z_1 indicates an approximate frequency of a pair of MMP zeros in *G(s)*.Similarly, Z_2 indicates another pair of MMP zeros. Consequently, the

sequence of the dominant modes in Fig 4-4a is "mode R → mode u → mode R → mode v", which is denoted as [R]-[u]-[R]-[v], based on our previous work (see Chapter 3). Since the phase of *G(s)* is the phase of the dominant mode, the phase of *G(s)* starts at −180° as mode R, drops to −360° due to the resonance of mode u, rises to −180° as mode R, moves to 0° as mode v, and ultimately drops to −180° due to the resonance of mode v.

As the variable ψ varies towards the curve veering point (i.e., $\psi=\psi_{cv}$), the spacing between mode u and mode v reduces. As a result, the combined contribution of these two modes in the frequency range (ω_u, ω_v) increases, which causes Z_1 and Z_2 to move towards each other. Since the peaks at ω_u and ω_v are infinite due to zero damping (Assumption A4-2), this combined contribution increases dramatically when mode u and mode v are sufficiently close. Therefore, there always exists an instance when Z_1 and Z_2 collocate and disappear, which is demonstrated in Fig 4-4b. As a result, the sequence of the dominant mode becomes [R]-[u]-[v]. Based on the mapping between the sequence of the dominant mode and the type of zeros (see Chapter 3), this sequence indicates the appearance of either CMP-CNMP zeros or two pairs of RMP-RNMP zeros (termed as dual RMP-RNMP zeros). Therefore, the phase of *G(s)* starts at −180° as mode R, drops to −360° due to the resonance of mode u, and ultimately drops to −540° due to the resonance of mode v. Moreover, for the dual RMP-RNMP case, in the evolution from dual MMP zeros to dual RMP-RNMP zeros, there must exist an instance when the zeros are CMP-CNMP zeros. Therefore, there exists a value $\psi=\psi_x$ such that CMP-CNMP zeros appear.

It is noteworthy that Condition C4-4y or the phase plot of *G(s)* in Fig 4-4a is commonly identified in flexible systems. This is because that the "rigid-body" mode (i.e., the double integrator or mode R in this chapter) commonly contributes a significant portion in *G(s)*. If there are two closely spaced modes that form a sequence of alternating signs together with the "rigid-body" mode, then there will likely be two pairs of MMP zeros between the two closely spaced modes. Examples will be provided in Section 4.3.

4.2.4. The Signs of the Modal Residues and the Sensor/Actuator Configuration Vectors

Based on the previous discussion, one can see that the signs of the modal residues play an important role in determining the types of zeros. This section provides some insights in understanding the signs of the modal residues in the view of the sensor/actuator configuration.

The eigenvalue problem of a system is in the form of $(K-\omega_i^2 M)\phi_i=0$, where M and K are the mass and stiffness matrices respectively, ω_i is the natural frequency of the ith mode, and ϕ_i is the corresponding eigenvector (mode shape) with its length normalized to unity (i.e., $\phi_i^T \cdot \phi_i=1$, for all i). Consequently, the modal residual of the ith mode is described as Eq. (4-15).

$$\alpha_i = \left(\beta_j^T \phi_i\right) \cdot \left(\beta_k^T \phi_i\right) / \left(\phi_i^T M \phi_i\right) = \left(\beta_j^T \phi_i\right) \cdot \left(\beta_k^T \phi_i\right) / \tilde{m}_i \tag{4-15}$$

In Eq. (4-15), $\tilde{m}_i$ is termed as the modal mass of the ith mode. Additionally, β_j and β_k are vectors in the same dimensions as ϕ_i. These two vectors define the sensor and actuator configuration of a system. For example, $\beta_j=[1\ 0\ \ldots\ 0]^T$ indicates that the sensor (or the actuator) is located on the point mass referring to the first coordinate. Since modal mass is always positive, the requirement that mode u and mode v are out-of-phase is then explained as Eq. (4-16).

$$\left[\left(\beta_j^T \phi_u\right)\left(\beta_k^T \phi_u\right)\right] \cdot \left[\left(\beta_j^T \phi_v\right)\left(\beta_k^T \phi_v\right)\right] < 0 \tag{4-16}$$

Next, the geometrical explanation of Eq. (4-16) is then presented. Fig 4-5a is a projection of the two eigenvectors ϕ_u and ϕ_v onto a 2D plane. The projection is conducted in a way such that the angle between ϕ_u and ϕ_v is unchanged on the 2D plane (i.e., $\cos\theta=\phi_u^T \cdot \phi_v$, where θ is the angle in the 2D plane). One can then draw two lines L(ϕ_u) and L(ϕ_v) that are perpendicular to ϕ_u and ϕ_v, respectively. As a result, the 2D plane is split into four quadrants (i.e., I, II, III, and IV). Therefore, Eq. (4-16) indicates that if one projects β_j and β_k onto this 2D plane, the two resulting vectors should be separately located in any two quadrants that are adjacent to each other (e.g., quadrants I and II in Fig 4-5a). It is noteworthy that due to the existence of curve veering, ϕ_u and ϕ_v are almost perpendicular to each other, since they are localized and the product $\phi_u^T \cdot \phi_v$ is almost zero (strict perpendicularity is reached in the case of mode crossing, since ϕ_u and ϕ_v are strictly localized). Moreover, as curve veering proceeds, ϕ_u and ϕ_v rotate with their relative angle

unchanged [95, 96], thus rotating the quadrants accordingly. As a result, this rotation may change the quadrants where the projected β_j and β_k are located, and may thus change the signs of mode u and mode v.

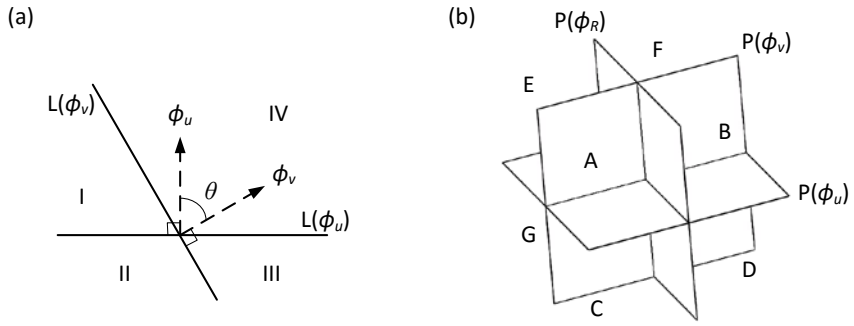

Fig 4-5 The spaces divided by the eigenvectors. (a) two dimensional quadrants; (b) three dimensional octants.

The above-discussed results can then be extended to a higher dimensional system. For example, consider a system under Assumption A4-1 to A4-5 (i.e., two 2nd order modes with a double integrator). Mode R, mode u, and mode v are the three modes of interest, with their natural frequencies ($0 \approx \omega_R < \omega_u < \omega_v$). If this system has a sequence of alternating signs (i.e., $\alpha_R \cdot \alpha_u < 0$ and $\alpha_u \cdot \alpha_v < 0$), the requirement on the sensor/actuator configuration vectors (i.e., β_j and β_k) is visualized in Fig 4-5b. The eigenvectors of ϕ_R, ϕ_u and ϕ_v can be projected onto a 3D space (not shown in Fig 4-5b for clarity). There are three planes (i.e., P(ϕ_R), P(ϕ_u), and P(ϕ_v)) shown in Fig 4-5b that are perpendicular to ϕ_R, ϕ_u, and ϕ_v, respectively. The 3D space is then split by these three planes into eight octants. The condition $\alpha_R \cdot \alpha_u < 0$ requires that the projection of β_j and β_k onto this 3D space should be separately located in two adjacent spaces divided by P(ϕ_R) and P(ϕ_u), such as A-B, A-C, or A-F; similarly, the condition $\alpha_u \cdot \alpha_v < 0$ requires that the projection of β_j and β_k should be separately located in two adjacent spaces divided by P(ϕ_u) and P(ϕ_v), such as A-E, A-C, or A-F. Therefore, for the conditions of both $\alpha_R \cdot \alpha_u < 0$ and $\alpha_u \cdot \alpha_v < 0$, the projection of β_j and β_k should be located in spaces meeting the requirements of both conditions. For example, if the projection of β_j onto this space is located in octants A, then the projection of β_k has to be located in octants C or octants F.

4.4. Case Studies

In this section, three examples are presented to demonstrate the correlation between curve veering and the presence of CMP-CNMP zeros. All of the examples follow the Assumption A4-1 to A4-3 (i.e., all of the decomposed modes are 2nd order systems, no unstable poles, and zero damping). The actuators considered in these examples are force actuators, while the sensors are displacement sensors.

4.4.1. Three DoF Spring-Mass Model

Consider the model shown in Fig 4-1a, using a numerical setting as $k_1=k_2=k_b=1$ kg/s^2, $m_b=10$ kg, $m_u=1$ kg, and $m_v=1+0.01$ kg. For the variation of parameter, we set k varying from 0 to 0.03 kg/s^2. The presence of curve veering is shown in Fig 4-1b. Next, using the coordinates as $[x_u, x_v, x_b]^T$, we consider the sensor/actuator configuration vectors as $\beta_j=[0.1148, 0.0733, 0.6873]^T$, $\beta_k=[0.1170, 0.2692, 0.5606]^T$.

The Bode plot of the relevant transfer function is shown in Fig 4-6a, with $k=0$ and $k=0.02$ kg/s^2 (which is close to the curve veering point, see Fig 4-1b). Mode u and mode v are the two modes dominated by the oscillations of m_u and m_v, while mode R is "rigid-body" mode where all of the three masses oscillate with almost the same magnitudes. In Fig 4-6a, one can see that mode R is well-separated from mode u and mode v. Therefore, in the vicinity of mode u and mode v, mode R can be regarded as a double integrator. Next, the poles and zeros related to the modes of interest are shown in Fig 4-6b. Note that the full pole-zero map is symmetric w.r.t. the real axis. It can be seen that the dual MMP zeros are observed in the range (ω_u, ω_v) when $k=0$. Therefore, based on Proposition P4-2y, there exists an instance of ψ (which is k in this example) when the dual MMP zeros become CMP-CNMP zeros (shown in Fig 4-6 when $k=0.02$ kg/s^2).

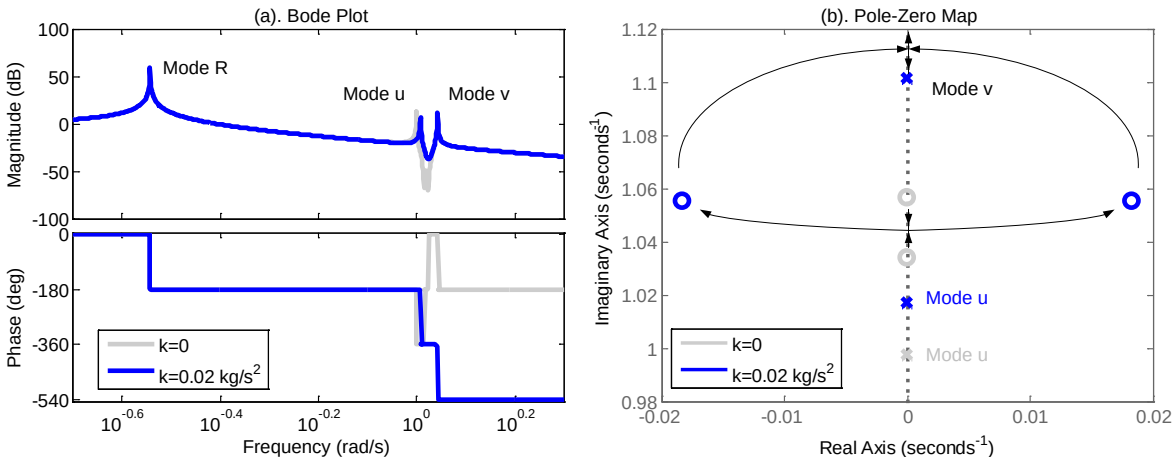

Fig 4-6 Dynamic properties with different values of k. (a) Bode plot; (b) pole-zero map.

The migration of zeros from k=0 to k=0.02 kg/s^2 is shown in Fig 4-6b. One can see that as k approaches the curve veering point, the spacing between ω_u and ω_v reduces (in this example, mode u moves towards mode v, see Fig 4-1b). As a result, as shown in Fig 4-6b, the two pairs of zeros merge towards each other, collocate, and then split into the complex plane, becoming a quartet of zeros symmetric w.r.t. to both the real and the imaginary axes (i.e., CMP-CNMP zeros).

In addition, it is noteworthy that if one continues to increase k, the relevant zeros will move back onto the imaginary axis again, but with a frequency higher than mode v, as shown in Fig 4-6.

4.4.2. Lumped Parameter Model of a Flexure Mechanism (SR4DoF)

Shown as Fig 4-7a, this model is from Chapter 2, which comprises two nominally symmetric double parallelogram flexure modules (DPFM). This mechanism allows stage ① to move in the X and Y directions. The former is due to the X direction bearings at stages ④ and ⑤ (i.e., rollers in Fig 4-7a) and the latter is due to large bending deformation of the beams in the two DPFM. The operating point (Y_{1o}) is given by a static displacement of stage ① in the Y direction

w.r.t. the ground, which is caused by a constant force F_{Y1o}. The five stages have eight displacement coordinates as shown in Fig 4-7a; lower case version of these coordinates represent the respective deviations w.r.t. to an operating point. Furthermore, each parallelogram (two for each DPFM) provides a degree-of-constraint. Therefore, this mechanism has four DoF.

There are two closely spaced modes (denoted as mode u and mode v) associated the oscillation of stages ② and ③. Based on Section 2.2.1, the two closely spaced modes are observable if there is certain asymmetric between the two DPFM. For simplicity, we define Δm_{23} ($=m_2/m_3-1$) as the mass asymmetric value. In this example, we consider Δm_{23}=1%, other numerical parameters are chosen based on Section 2.2.1. The natural frequencies of these two closely spaced modes are examined in Fig 4-7b, using the operating point as the variation parameter. One can see that curve veering is observed (mode crossing is achieved when Δm_{23}=0).

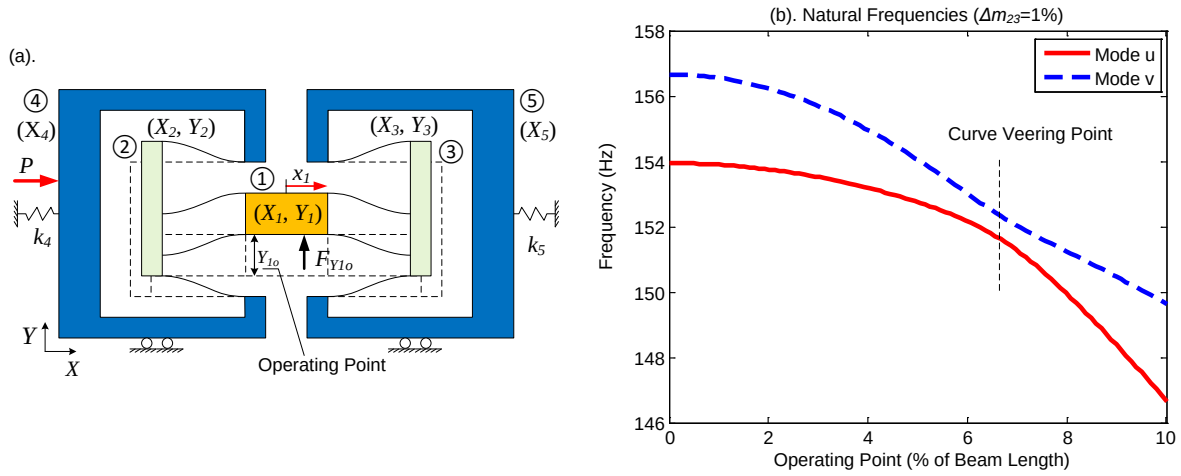

Fig 4-7 Lumped parameter flexure model. (a) schematic diagram; (b) natural frequencies of the closely spaced modes.

The coordinate chosen is $[x_1, y_1, y_2, y_3]^T$. The non-collocated transfer function from an X direction force P on stage ④ to the X displacement of stage ① is investigated, under different values of Y_{1o}. As a result, the sensor/actuator configuration vectors are computed as $\beta_j=[1, \alpha, -2\alpha, 0]^T$ and $\beta_k=[1, 0, 0, 0]^T$, where α is a variable defined as $\alpha=0.6\cdot Y_{1o}/L$ and L is the length of the beams in each DPFM (see Section 2.2.1).

The Bode plot of the relevant transfer function is shown in Fig 4-8a, with Fig 4-8b as a zoom-in view of the modes of interest. Among the four modes of this model, there is one mode that is associated with the oscillations of stages ①, ②, and ③ in the Y direction, making it unobservable in the X direction transfer function. Next, one can see that mode R is well-separated from mode u and mode v, and dual MMP zeros are observed when Y_{1o}=3.3% of L. Therefore, based on Proposition P4-2y, CMP-CNMP zeros are observed in Fig 4-8, as Y_{1o} increases towards the curve veering point.

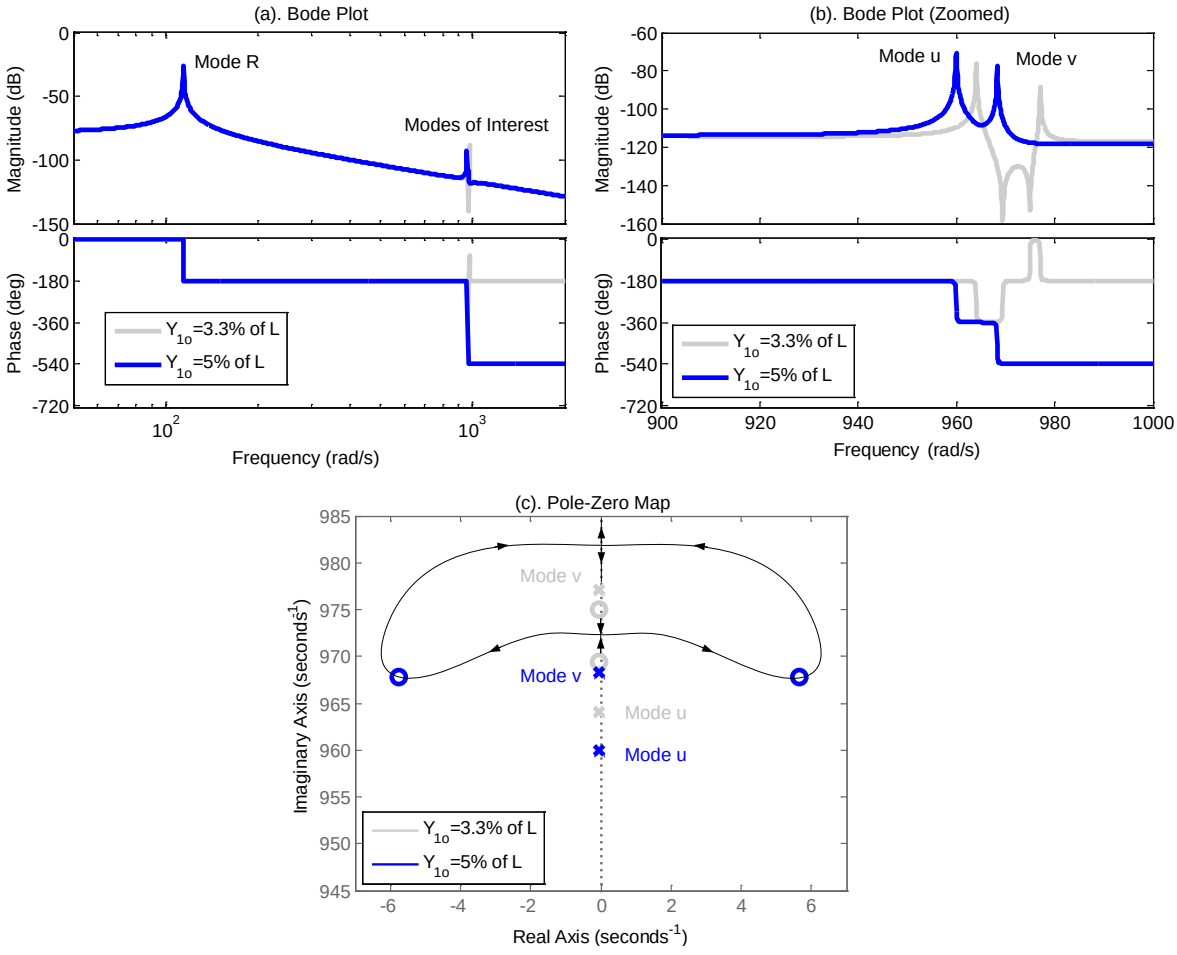

Fig 4-8 Dynamic properties of the SR4DoF flexure mechanism. (a) Bode plot; (b) zoom-in Bode plot; (c) pole-zero map

The migration of zeros as Y_{1o} varies is shown in Fig 4-8c. Similar to Fig 4-6b, it can be seen that the zeros starts as dual MMP zeros, merge towards each other, collocate, split into the complex plane (CMP-CNMP zeros), and move back on to imaginary axis (dual MMP zeros).

4.4.3. Four DoF Mass-Spring Model

Consider the lumped parameter model shown in Fig 4-9a. This example is inspired by the work of Loix et al. [49], who conducted a numerical study of a 4 DoF mass-spring model and observed CMP-CNMP zeros. However, they did not realize the presence of curve veering and the correlation between CMP-CNMP zeros and curve veering.

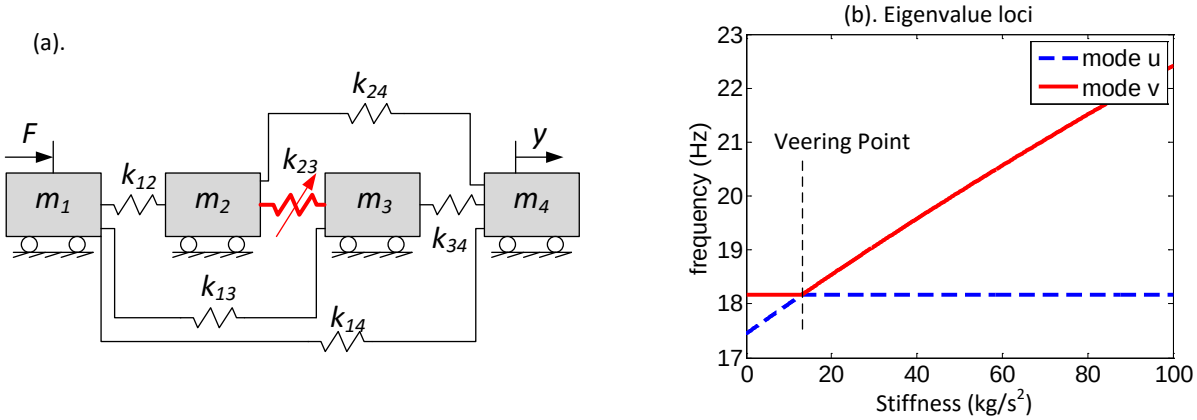

Fig 4-9 Four DoF mass-spring model. (a) schematic diagram; (b) natural frequencies of the two modes of interest.

Define the coordinates as $[x_1, x_2, x_3, x_4]^T$ and consider the dynamics. The numerical values used in this example are set as m_1=m_4=10 kg, m_2=m_3=1 kg, k_{12}=k_{34}=100 kg/s^2, and k_{13}=k_{14}=k_{24}=200 kg/s^2. k_{23} is the parameter for variation. One can see that m_1 and m_4 are set with values much larger than the other two, in order to explicitly demonstrate closely spaced modes. The two closely spaced modes are dominated by the oscillation of m_2 and m_3, which are again denoted as mode u and mode v. The curve veering is shown in Fig 4-9b. Next, regarding the transfer function, the sensor/actuator configuration vectors are β_j=$[1, 0, 0, 0]^T$ and β_k=$[0, 0, 0, 1]^T$, which indicates that the sensor and the actuator are located on the lumped mass.

The Bode plot of the relevant transfer function is shown in Fig 4-10a. Since there are four modes (including a double integrator), one cannot easily determine whether Assumptions A4-4 and A4-5 apply to this example. However, it can be seen that dual MMP zeros are observed between the two closely spaced modes. Therefore, the presence of the dual MMP zeros implies that the variation of zeros follows a procedure similar to what is described in Fig 4-4. In other words, as the spacing between mode u and mode v reduces due to curve veering, the contribution of mode u and mode v in the range (ω_u, ω_v) is "squeezed". As a result, the dual MMP zeros merge towards each other and split into the complex plane, becoming CMP-CNMP zeros.

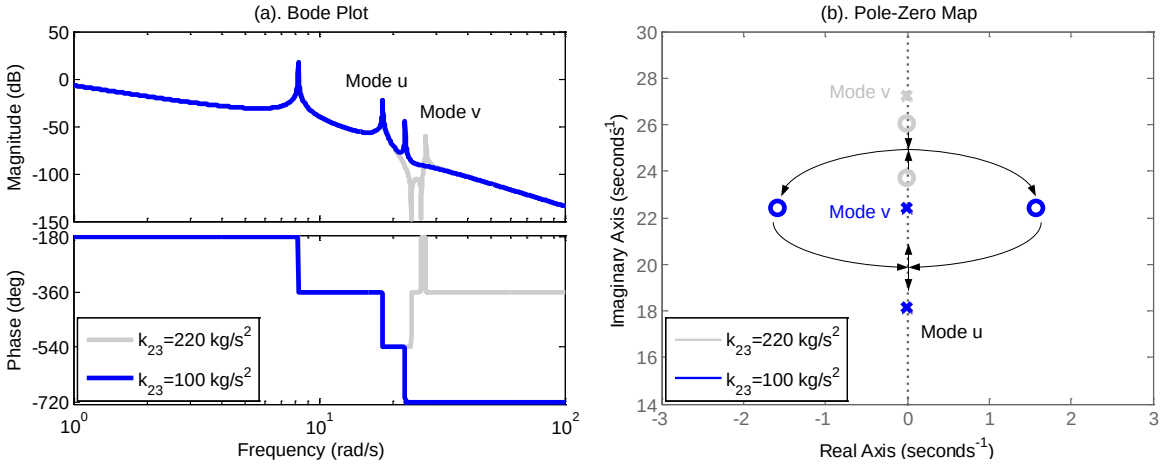

Fig 4-10 Dynamic properties of the 4DoF model. (a) Bode plot; (b) pole-zero map.

The migration of zeros is demonstrated in Fig 4-10b. It is noteworthy that this migration is similar to what are described in Fig 4-6b and Fig 4-8c. However, Proposition P4-2y only explains part of the procedure (i.e., merge, collocate, and split) that is related to curve veering, while the explanation of the full procedure is currently open and being investigated.

4.5. Contributions and Conclusions

Three main findings are presented in this chapter. First, we demonstrate that under certain sensor/actuator configurations, the presence of curve veering in a system guarantees the presence

of CMP-CNMP zeros (sufficient condition). However, the presence of CMP-CNMP zeros does not guarantee the presence of curve veering (not a necessary condition). We show an example where CMP-CNMP zeros appear but there is no curve veering. Second, based on the importance of the signs of the decomposed modes, we show the relationship between sensor/actuator configuration and the signs of the decomposed modes. This finding provides a useful insight in designing the sensor/actuator of a system. Third, three examples are provided showing the correlation between curve veering and CMP-CNMP zeros.

Chapter 5. Experimental Validation of CNMP Zeros in Flexure Mechanisms

In this chapter, we present an experimental validation on the existence of CMP-CNMP zeros in lightly damped, passive stable flexure mechanisms (i.e., there is no right half plane poles of the plant). Also, this work validates the model-based prediction of CMP-CNMP zeros in double parallelogram flexure module based mechanisms as in Chapter 2. Experimental validation of the model is challenging because it requires building an experimental setup that captures the dynamics of the theoretical model. However, the dynamic behavior of an actual setup can be adversely affected by the several factors that have to be carefully considered, as listed below.

1. The SR4DoF model is designed for in-plane translational motion. Since there are always un-modeled modes in reality, it follows that the un-modeled modes have to be designed at frequencies much higher than the modeled modes, so that they have minor contribution (to magnitude or phase) in the frequency response at low frequencies (i.e. the frequency range of interest). These un-modeled modes include out-of-plane modes, in-plane rotational modes, or additional distortion modes due to the compliance of the "supposedly rigid" stages of the flexure mechanism.
2. The CNMP zeros of the SR4DoF model are the results of the superposition of two closely spaced modes (i.e., modes of interest) and a non-zero but relatively small contribution of all other modes. Therefore, these modes of interest have to be designed to be well-separated from any "rigid-body" mode at the lower frequencies and the un-modeled modes in the higher frequencies.
3. The CNMP zeros of the SR4DoF model are predicted at various operating points. The method to provide a quasi-static actuation to establish the operating point has to be such that this actuator minimally affects the dynamics of the flexure mechanism. Providing such actuation over a range of displacements in a perpendicular direction can be challenging for the setup design.
4. All bearings have to be designed to avoid nonlinearities such as friction and backlash.

To conduct the validation, the design procedure is presented to minimize the extraneous factors in order to provide a "clean" observation. First, iterations are discussed on the geometric design of the experimental setup. Next, dimensions are optimized in terms of mode orders and natural frequencies, in order to isolate the modes of interest from other modeled and not-modeled modes. Then, fabrication of this setup is discussed. Then, frequency responses are experimentally measured and compiled over a two-variable based grid, which agrees with a model based prediction of the existence of CMP-CNMP zeros (i.e., the CNMP map). Via the validation of this map, we show that an intentionally asymmetric choice of mass can help eliminate the CMP-CNMP zeros.

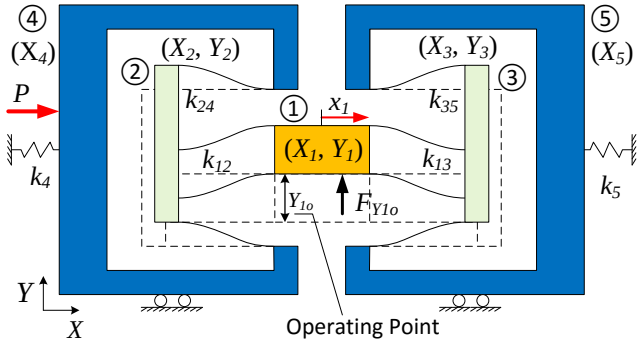

Fig 5-1 Simple representative 4 DoF model (recall of chapter 2).

The SR4DoF model is shown again in Fig 5-1 for convenience. The closely spaced modes of interest are the modes dominated by stage ② and ③. We have shown in Chapter 2 that these closely spaced modes interact as the above mentioned coupling varies with the operating point and as the parametric asymmetry varies, giving rise to CMP-CNMP zeros under certain conditions. This is again shown in Fig 5-2 for convenience, which maps the existence of CMP-CNMP zeros against a range of operating points (Y_{1o} normalized with respect to beam length L) and parametric asymmetry (Δm_{23}). Black regions on this map indicate the presence of CMP-CNMP zeros. For ease of presentation, the top view (Fig. 2a) is referred to the CNMP zero map for ease of presentation.

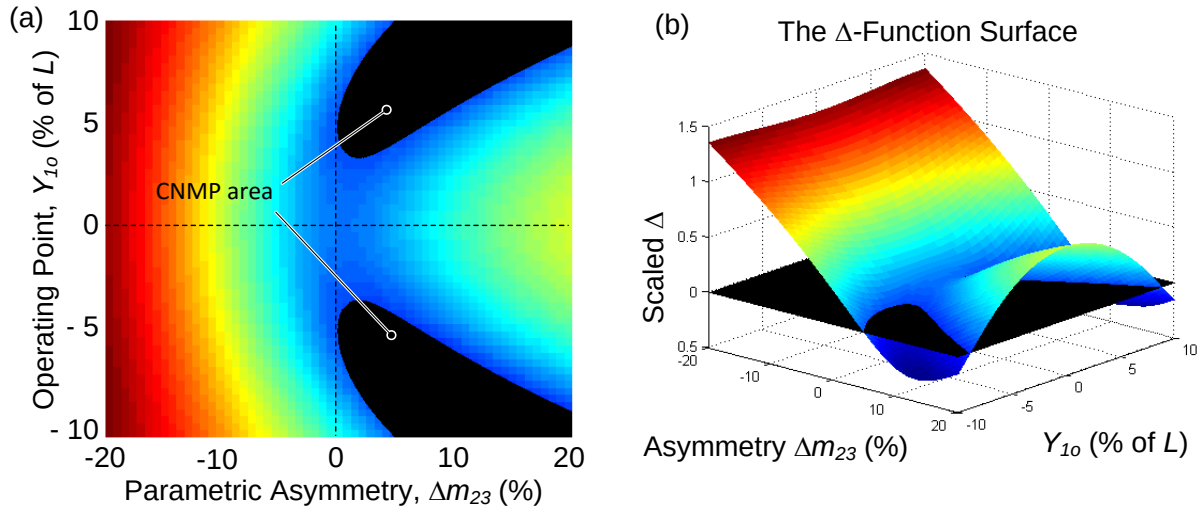

Fig 5-2 Complex non-minimum phase zero map (recall of chapter 2).

Two main observations are presented on the CNMP map (Fig 5-2a). First, the map is symmetric w.r.t. the operating point (i.e., Y_{1o}). This is due to the fact that different signs of the operating point indicate that the relevant beams bend in different directions. Since the same amount of bending in different directions results in the same amount and same sign of kinematic error motion and the CNMP behavior is related to such kinematic error motion, the discriminant function is symmetric w.r.t. the operating point.

Second, the CNMP phenomenon is directional w.r.t. the parametric asymmetry (i.e., Δm_{23}). In other words, when $\Delta m_{23}>0$, the CMP-CNMP zeros appear under certain conditions; however, when $\Delta m_{23}<0$, the CMP-CNMP zeros will not appear. A primary explanation for this directional behavior is that the non-collocated actuator/sensor configuration of this system is directional (i.e., from the force applied on stage ④ to the displacement sensor on stage ①). The study in Chapter 3 shows that the sign of Δm_{23} is related to the sequence of signs of the decomposed modes, which is shown to be relevant to the existence of CMP-CNMP zeros.

However, the above results and predictions are based on numerical simulations of a model with various assumptions. Thus, an experimental validation of this model is of great value to corroborate the understanding of the phenomenon related to CMP-CNMP zeros.

5.1. Experimental Setup Design

5.1.1. Flexure Mechanism Geometry

The goal is to design an experimental setup whose low-frequency modes (i.e., main dynamics behavior) are the modes of the SR4DoF model, while the un-modeled modes show up at high frequencies. As the first step, in order to minimize other extraneous factors such as friction and backlash, the roller bearings are replaced with flexure bearing. This enables the setup to be manufactured from a flat metal plate. Next, Aluminum 6061-T651 (25.4 mm thick stock plate) was selected as the primary option for ease of availability and machining. In addition, voice coil motor was chosen for the actuation of the X-axis motion, since it provides a non-contact actuation between the ground and the flexure mechanism.

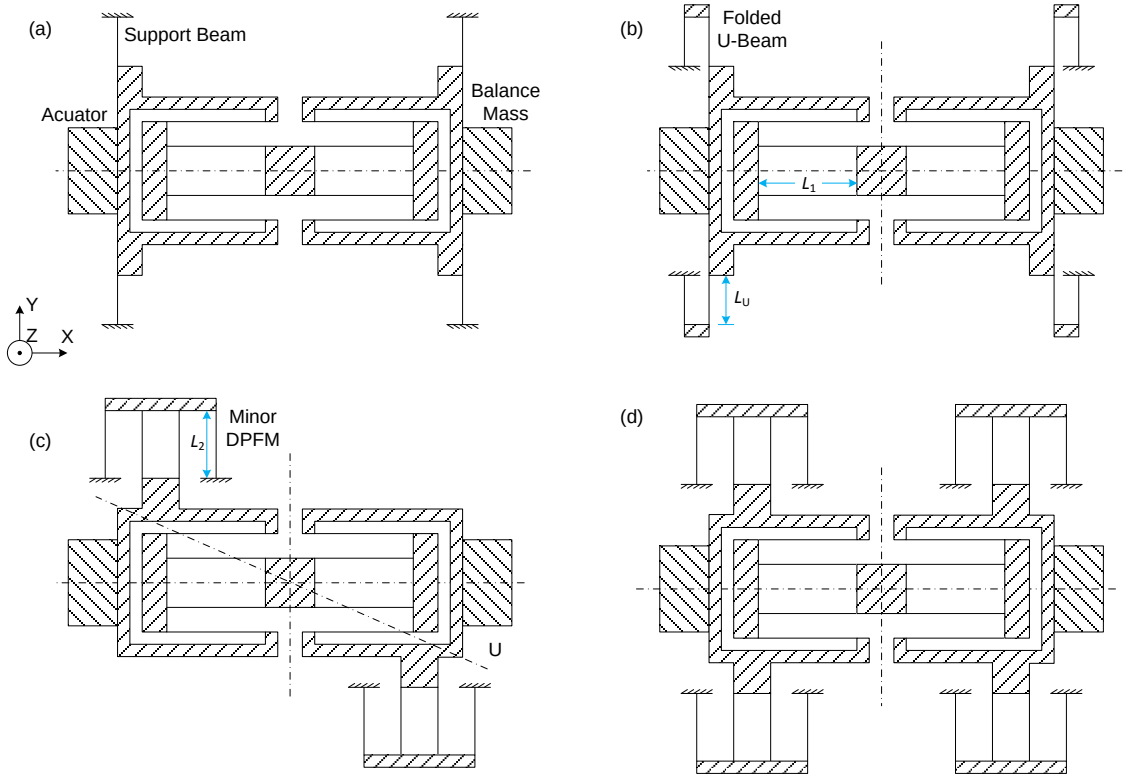

Fig 5-3 Illustration of geometric evolution.

The initial design shown in Fig 5-3a (namely Design-1) utilizes four simple flexure beams to provide flexure bearing for stages ④ and ⑤. A balance mass with the same mass value of the

mover of the actuator (i.e., coil) is assigned on the opposite side of the actuator to achieve a nominally mass balance between for stages ④ and ⑤. The reason to pursue nominally mass balance is to increase the frequencies of several rotational modes, as large mass imbalance could lead to dramatic increase in the momentum w.r.t. certain out-of-plane rotation (e.g., pitch) and thus reduce the natural frequency of the associated un-modeled mode. However, using the simple beams has limited travel range in the X direction due to geometric over-constraint of these support beams.

To relieve the over-constraint in the design shown as Fig 5-3a, the simple beams are replaced by folded beams as shown in Fig 5-3b, namely Design-2. To enable the finite element analysis (FEA) of the modes shapes and their corresponding orders, we established the dimensions of the main flexure beams that were used in our previous investigations [11, 72]. Therefore, as the initial trial, the length of the main beam was set at 47.5 mm (i.e., L_1 in Fig. 3b) and thickness was set at 0.625 mm (denoted as T_1). Regarding the folded U-beam, the length was set at 40 mm (i.e., L_U in Fig 5-3b) and the thickness was set at 0.625 mm. In additional, the masses were designed by setting m_1=0.2 kg, m_2=m_3=0.04kg, and m_4=m_5=0.8 kg (including the mass value of the actuator/balance mass). The purpose of employing such parameter choice was to create a desktop size experimental flexure mechanism.

FEA was conducted on Design-2, with stage ① at the nominal operating position (i.e., Y_{1o}=0). The nonlinear option of the FEA was turned off; even though our previous studies show that the natural frequencies of the two closely spaced modes will have about 7% changes under large deformation [72]. The reason is that, the primary objective of using FEA in this study is to arrange the un-modeled modes to high frequencies, while the change of the modeled modes' frequencies has minor priority.

The FEA results of the Design-2 are presented in Table 5-1. The modes in shaded color are the un-modeled modes of the abstract model. It can be seen that the first two modes are already modeled modes. However, there are several un-modeled modes appearing between the second and third modeled modes (i.e., the 2nd and 8th modes). For example, the 3rd mode is the in-plane rotational mode as shown in Fig. 4a, while the 7th mode is the out-of-plane rotational mode as shown in Fig 5-4b. The reason is that the folded U-beam has low stiffness in both the translation

and rotation in the axial direction [97]. However these directions are the bearing directions of the SR4DoF model.

Table 5-1 FEA results of each design. Shaded areas refer to the un-modeled modes. The unit of frequency is Hz.

Order	Design-2 (Fig. 5-3b)		Design-3 (Fig. 5-3c)		Design-4 (Fig. 5-3d)		Design-4x (Improved)	
	Freq.	Description	Freq.	Description	Freq.	Description	Freq.	Description
1	15.2	"Rigid-body"	10.6	"Rigid-body"	14.8	"Rigid-body"	25.2	Vertical[a]
2	35.5	Vertical[a]	29.7	Vertical[a]	29.7	Vertical[a]	33.1	"Rigid-body"
3	43.8	In-plane	44.9	Out-of-plane	83.0	CSM1[b]	79.6	CSM1[b]
4	46.5	In-plane	51.5	Out-of-plane	91.3	CSM2[b]	85.6	CSM2[b]
5	54.1	In-plane	82.3	CSM1[b]	123.8	Out-of-plane	235.1	Out-of-plane
6	57.0	Out-of-plane	91.3	CSM2[b]	128.0	Out-of-plane	237.8	In-plane
7	67.0	Out-of-plane	111.2	In-plane	143.6	In-plane	272.8	In-plane
8	103.4	CSM1[b]	112.3	In-plane	144.2	In-plane	281.4	Out-of-plane
9	116.1	CSM2[b]	149.6	In-plane	144.8	In-plane	282.0	In-plane
10	187.5	Out-of-plane	159.1	Out-of-plane	145.6	In-plane	283.6	In-plane

Note:

[a]: Vertical mode is the modeled mode where stages ①, ② and ③ oscillate vertically in the Y-axis.

[b]: "CSM" refers to the two closely spaced modes dominated by the Y-axis oscillation of stages ② and ③. CSM1 refers to the mode where stages ② and ③ oscillate in the opposite directions, while CSM2 refers to the case where stages ② and ③ oscillate in the same direction.

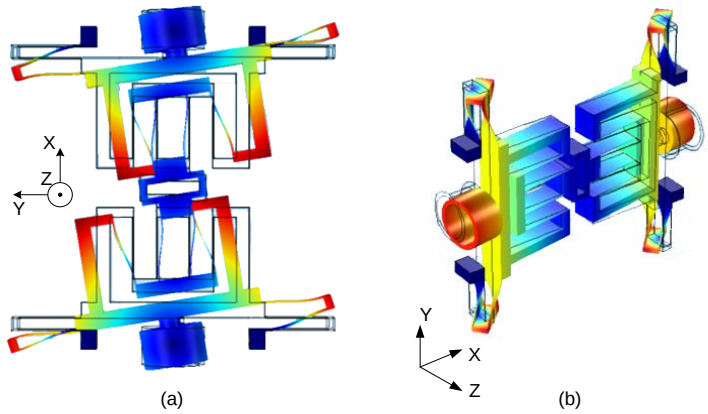

Fig 5-4 Examples of un-modeled modes in Design-2: (a) 3rd mode; (b) 7th mode.

To increase the stiffness in the relevant bearing directions, the folded U-beam structures in Fig 5-3b are replaced with two additional DPFMs as shown in Fig 5-3c (denoted as Design-3). The reason is that DPFM offer better stiffness characteristics without sacrificing the travel range

[30, 97]. The additional DPFMs are denoted as the minor DPFMs to distinguish from the two DPFMs (denoted as the main DPFMs) used in the SR4DoF model. The beams of the minor DPFMs were initially designed with the same dimensions as the beams in the main DPFM, with length (L_2) set as 47.5 mm and thickness (T_2) set as 0.625 mm. In addition, in order to reduce the frequencies of the two modeled closely spaced modes (i.e., CSM1 and CSM2 in Table. 1), m_1 was increased from 0.2 kg to 0.4 kg, and m_2 and m_3 were increased from 0.04 kg to 0.06 kg. The re-arrangement in mass value was to create a larger frequency separation between the last modeled mode (i.e., CSM2) and the next un-modeled mode.

The FEA results are shown in Table 5-1. It can be seen that the number of un-modeled modes between the modeled modes have reduced. However, since the geometry of Design-3 is not symmetric w.r.t. both the X and Y axes, the flexure mechanism has an axis connecting the grounds (denoted as U-axis in Fig 5-3c) with low rotational stiffness. In fact, the un-modeled out-of-plane modes shown as the 3rd and 4th modes in Design-3 are associated with the rotation w.r.t. the U-axis.

An intuitive improvement to Design-3 was to employ four minor DPFMs to achieve symmetric w.r.t. both the X and Y axes, resulting to Design-4 as shown in Fig 5-3d. It can be seen from the FEA results (Table 5-1) that the rotational modes (w.r.t. the U-axis in Fig 5-3c) were effectively eliminated. Moreover, all of the first fourth modes were modeled modes. However, frequency separation between the last modeled mode (91.3 Hz) and the first un-modeled mode (123.8 Hz) has a factor of only 1.35 (123.8/91.3). Moreover, the four closely spaced modes (from 7th to 10th) are associated with the oscillation of the secondary stages in the four minor DPFMs. These closely spaced modes also need to be re-arranged to higher frequencies. Therefore, further considerations are needed to increase all these un-modeled modes' natural frequencies.

The 5th and 6th modes are shown in Fig 5-5a and Fig 5-5b, respectively. For both the 5th and 6th modes, the center stage (stage ①) is almost stationary, thus dividing the flexure mechanism into two parts symmetric w.r.t. the Y-axis that is located on the center stage. Furthermore, each part rotates w.r.t. the Y-axis that is located on the part that connects the minor DPFM and the main DPFM (denoted as W-axis, parallel to the Y-axis). The relevant angular momentum of such rotations is mainly determined by L_3, L_4, and L_5 as shown in Fig 5-5d.

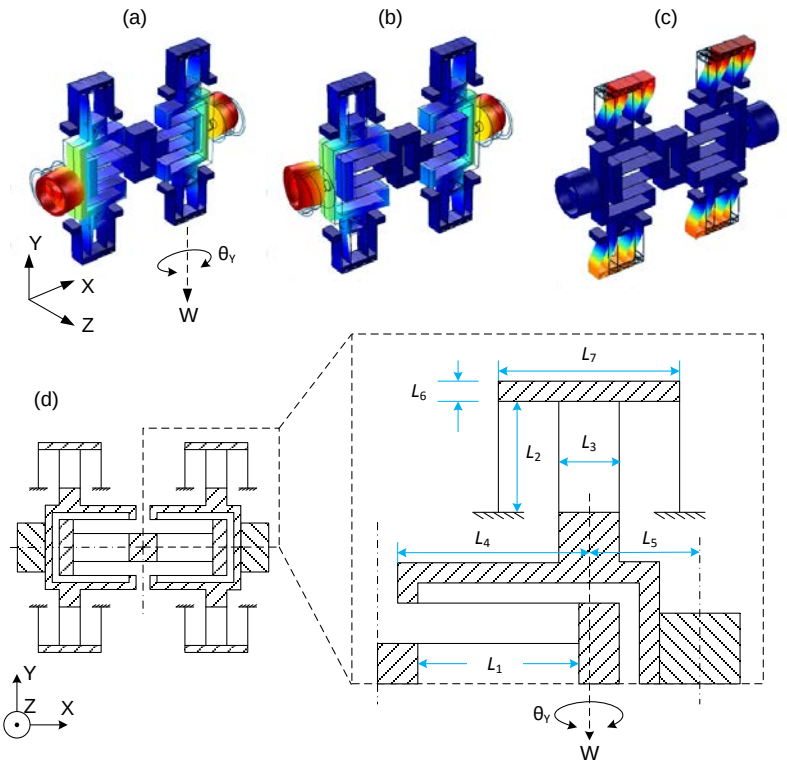

Fig 5-5 Examples of the un-modeled modes in Design-4 and the considerations on dimensions: (a) 5th mode; (b) 6th mode; (c) 7th mode; (d) dimensions of Design-4 under consideration.

The consideration on L_3, L_4, and L_5 are shown as below.

1. L_3 is the span of the inner parallelogram of the minor DPFM, which affects the stiffness of rotation w.r.t. the W-axis. When L_3 is closed to zero, the rotational stiffness will be dominated by the torsional stiffness of the beam; when L_3 is increased, a new factor will contribute to the rotational stiffness, which is the Z-axis transverse stiffness of the beam and the span of the inner parallelogram. For a single beam with the typical dimension in this study (e.g., T=0.625 mm, H=25.4 mm, L=47.5 mm), the rotational stiffness due to torsion is 1.099 (N·m/rad) at the tip [98]; while rotational stiffness due to z-axis transverse stiffness is 3.714 (N·m/rad) with a lever arm of 1 mm (i.e., half of the span). Therefore, increasing L_3 will increase the rotational stiffness. However, increasing L_3 will lead to a larger size of secondary stage of the minor DPFM, which could reduce the natural frequencies of the modes associated with the minor DPFM. At this moment, L_3

was increased from 20 mm to 40 mm. The dimensions of the secondary stage of the minor DPFM (i.e., L_6 and L_7) are discussed later in this section.

2. L_4 and L_5 affect angular momentum of the associated rotational mode. Thus, the design objective is to have the mass center of stage ④ or (stage ⑤) lie in the W-axis, such that one can minimize the angular momentum w.r.t. the W-axis, thus increasing the natural frequencies of the corresponding rotational modes. At this moment, we chose L_4=85mm and L_5=30mm.

The 7th mode is shown in Fig 5-5c, as an example of the four closely spaced modes associated with the secondary stage of the minor DPFM. To increase the natural frequencies of these modes, one can either increase the transverse stiffness of the minor beams (related to the length L_2 and the thickness T_2), or reduce the mass of the minor secondary stages (related to the dimensions L_6 and L_7). However, there are tradeoffs in both approaches. The considerations are show as below

1. Increasing the transverse stiffness (X-axis) of the minor DPFM also increases the frequency of the "rigid-body" mode (i.e., the 1st mode in Design-4 as in Table 5-1). This tradeoff was carefully balanced to create a good frequency separation for the modes of interest (i.e., CSM1 and CSM2) from both the "rigid-body" mode and the un-modeled modes related to the minor DPFMs. By iteratively selecting dimensions and checking the mode shapes and frequencies. At this moment, we chose L_2=47.5 mm, T_2=1 mm.
2. Reducing L_7 results in a smaller span of the outer parallelogram of the minor DPFM, thus lowering the rotational stiffness in the Y-axis; reducing L_6 increases the compliance of the minor secondary stages. This leads to a lower axial stiffness and introduces additional un-modeled modes at lower frequencies. At this moment, we chose L_7=80 mm, which was twice as the inner span L_3. Since the lever arm affects the rotational stiffness in a quadratic manner, the rotational stiffness of the outer parallelogram was satisfactory and will not introduce additional un-modeled modes in low frequencies. In addition, we chose L_6=3 mm. The compliance due to small value of L_6 is discussed next.

The FEA results of the improved design (denoted as Design-4x) are shown in Table 5-1. The first eight modes are shown in Fig 5-6. Several observations are provided as below.

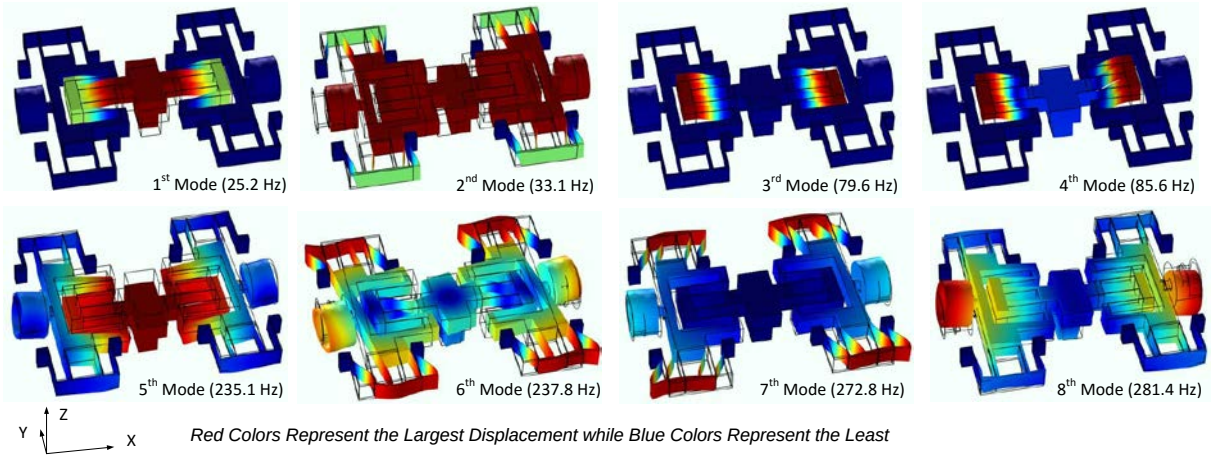

Fig 5-6 Mode shapes of Design-4x (final design).

1. It can be seen that the "rigid-body" mode appears as the 2nd mode due to increase in the transverse stiffness of the minor DPFMs. However, since the 1st mode is associated with the Y-axis oscillation of stages ①, ② and ③, it is nearly unobservable in the frequency response measurement at stage ① in the X-axis.
2. The 3rd and 4th modes are the modes associated with the oscillations of two secondary stages (stages ② and ③), which are the modes of interest in this chapter. They are closely spaced compared to their adjacent modes (the 3rd and 5th modes).
3. The 5th mode is an out-of-plane mode in the Z-axis due to the increase in the mass value of stage ①. The 6th mode is an in-plane rotational mode w.r.t. the Z-axis, due to the compliance of the secondary stages of the minor DPFMs. However, since the 5th and 6th modes appear at frequencies around 235 Hz with a factor of almost 3 from the 4th mode, their contributions to the low frequencies are minor.
4. The four closely spaced modes associated with the four minor DPFMs are shown around 280 Hz. One of these modes is the 7th mode shown in Fig 5-6. Comparing to Design-4, they were increased by a factor of two and had a factor of 3.5 from the 4th modeled mode.
5. The 8th mode in Design-4x is the same as the 5th mode in Design-4, whose frequency was increased from 123.8 Hz to 281.4 Hz. In addition, the 6th mode in Design-4 was shown as the 13th mode in Design-4x that is not shown in Fig 5-6.

In summary, in the final design, there is at least a two-times frequency separation before and after the closely spaced modes (3rd and 4th) of interest, offering a satisfactory isolation from the other modes. Therefore, the dimensions in Table 5-1 were employed in the fabrication of the final experimental setup hardware.

5.1.2. X-Axis Motion Actuation and Sensing

The primary experiment is to conduct system identification at different operating points (Y_{1o}). While the range of motion in the Y-axis is large (e.g., ±5 mm, referring to about 10% of the beam length), the range of motion for X-axis can be small. This provides convenience in the selection of actuator and sensor for the X-axis motion. Based on a factor of two for safety against yielding, we designed the range of motion for X-axis as ±1 mm. In addition, a mechanical stop was employed to prevent over-ranged displacement.

Accordingly, the actuation force required for 1 mm displacement is 72 (N). However, the experiment did not require the flexure mechanism to be held at 1 mm displacement in the X-axis. Instead, the system identification was conducted by sending an up-chirp sinusoid signal with zero mean. Therefore, the actuator need not to provide a continuous force of 72 (N). In our design, a voice coil motor (BEI Kimco Magnetics, model: LA24-20-000A, force constant: 11.12 N/A, stroke: ± 8.26 mm, peak force 111 N, continuous force: 26 N) was selected for the X direction actuation. An L-shape part was designed to provide connection between the flexure mechanism and the mover of the motor. The stator of the motor was directly mounted on the ground frame via two X-axis mounting holes. The concentricity tolerance between the mover and the stator was allowed by the air gap of the voice coil motor.

Regarding the sensing, non-contact sensors are ideal for this experiment since they do not have friction and backlash. Moreover, the center mass (stage ①) was designed to have displacements in both the X and Y axes. The sensor should be able to accommodate large displacements in the Y-axis. Capacitance probe (Lion Precision, range: 2 mm; peak-peak resolution: 400 nm; RMS resolution: 40 nm) was chosen to measure the X-axis displacement of

the center mass (stage ①) w.r.t. the ground. A bridge was designed to mount the capacitance probe. Both ends of this bridge were directly connected to the ground frame.

5.1.3. Constant Force Loading Mechanism

For the SR4DoF XY mechanism shown in Fig 5-1, the operating point is set by applying a static Y direction force F_0. One of the challenges in applying such a constant force on the motion stage is that the latter's position varies along the X and Y axes. During the frequency response measurement, the motion stage vibrates in the X direction with a small amplitude in response to an excitation in the X direction. Also, due to cross-axis coupling, there can be an even smaller but finite Y direction vibration at the motion state in response to an X direction excitation. The Y direction actuation system has to accommodate these X and Y direction displacements during a frequency response measurement. Furthermore, since this measurement has to be repeated at various Y direction operating points (Y_{1o}=±5mm), the Y direction actuation system has to be capable of providing the necessary force (F_0) over this range and hold it constant at a given operating point.

A potential option is to employ non-contact force actuator such as a voice coil motor or a moving magnetic actuator [99]. Apart from the cost consideration, two main challenges exist: first, to provide desirable Y-axis displacement, one needs to apply either closed loop position control or complex current management for the actuator at different Y-axis displacement. Second, it is difficult to guarantee a constant force when the motion stage oscillates in the X-axis with different frequencies, such as following an up-chirp sinusoid signal. The solution to these challenges will lead to unnecessary complexity of the setup design.

Instead of employing an electromagnetic actuator to apply a non-contact constant force, we chose a relatively simpler actuation method utilizing steel braided cables and free weights, using a passive actuation. However, this method requires the use of a pulley to change the direction of force provided by the weights. Since a traditional pulley introduces displacement uncertainty in the loading direction due to rolling and/or sliding friction, a virtual pulley was employed [3, 100].

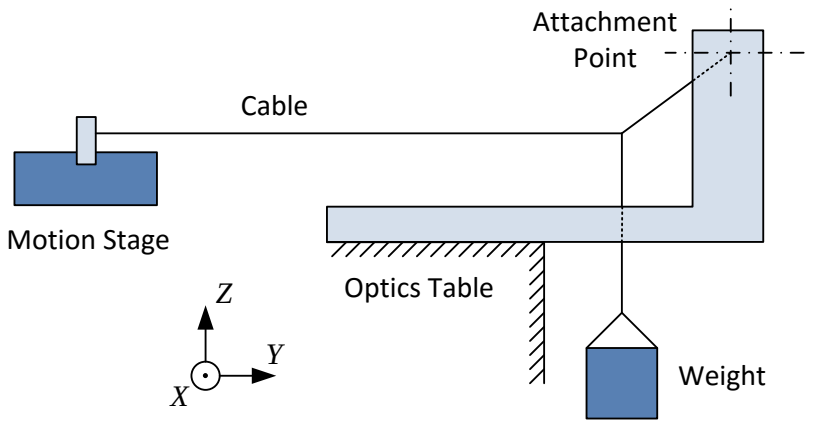

Fig 5-7 Virtual pulley with long cable.

Fig 5-7 shows the schematic of the virtual pulley arrangement comprising three cable segments: a long horizontal segment, a vertical segment, and a diagonal segment. All three segments are connected at a common point. The other end of the horizontal segment is attached the motion stage; the other end of the vertical segment is connected to a hanging weight; and the other end of the diagonal segment is connected to a rigid extension of the ground reference (e.g. optics table). This cable arrangement is called a virtual pulley because it converts the vertical force of the weight in the Z direction to a horizontal force F_0 on the motion stage in the Y direction. The two are simply related by the trigonometry of the cable arrangement. In this manner, a constant Y direction force F_0 is achieved by a known hanging weight, without the need for any active force control. The horizontal section of the cable was selected to be long enough so that it can absorb any X direction displacement of the motion stage without appreciably impacting the F_0 force. Any Y direction vibration of the motion stage is also easily accommodated by the virtual pulley. The Y direction operating point is varied simply by changing the weight, and keeping track of any change in the cable arrangement trigonometry. Thus, the virtual pulley provides a frictionless pulley functionality because it eliminates any rolling or sliding interfaces and any associated uncertainties

The limitation of this design is a relatively low accuracy in the measurement of the operating point, comparing to a position sensor such as linear encoder. Since the operating point is derived by the applied weight and the angles of the virtual pulley, the relatively bias of the angle measurement could lead to a scaling error in the derived operating point. However, since the

scaling error does not affect the distribution pattern of the CNMP map, the virtual pulley loading mechanism was employed for the experiment.

5.2. Experimental Results and Discussion

5.2.1. Hardware Fabrication and Assembly

The experimental setup hardware was fabricated and assembled for testing, as shown in Fig 5-8. The flexure mechanism, along with the reference ground frame, were made monolithically from a 25.4-mm thick AL6061-T651 plate using wire-electric discharge machining (EDM), which has a global dimensional tolerance of ± 0.005 mm. The other parts for assembly were fabricated by traditional machining, such as the coupling between the mover and the flexure, the bridge to mount the capacitance probe, and the target block for capacitance probe. Four standings were used to connect the ground frame of the flexure mechanism to an isolation table (Newport RS1000), in order to minimize the deformation of the ground frame.

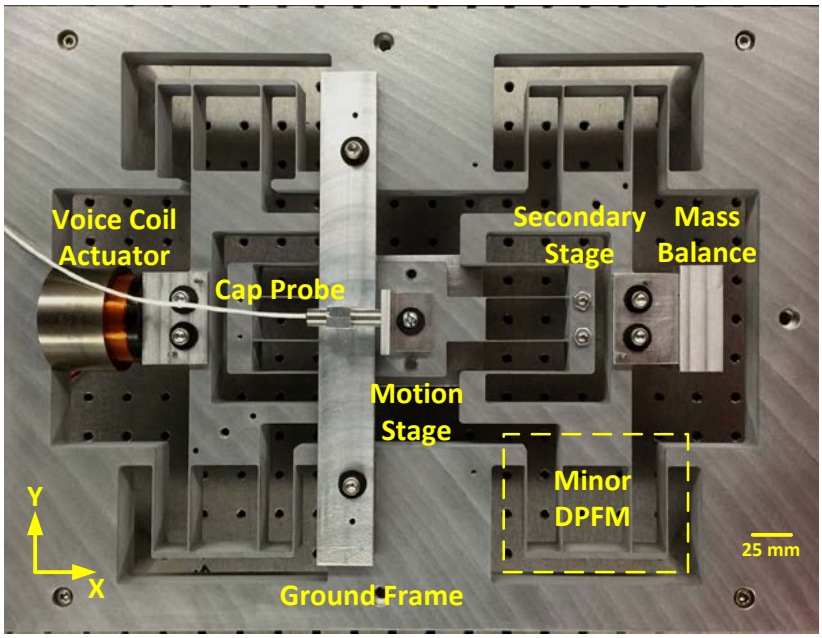

Fig 5-8 Experimental setup hardware.

Assembly holes were fabricated on the two secondary stages (stages ② and ③), as shown in Fig 5-8. The intended mass asymmetry between stages ② and ③ was achieved by adding incremental small masses with a value of 0.5 g, which is 0.8% of the secondary stage (60 g).

A custom-built current amplifier based on a Power OpAmp MP111 amplifier was used to drive the current in the voice coil actuator at a bandwidth of 1 kHz [1]. The 1 kHz bandwidth of the current loop is sufficient high to avoid interference in the dynamics of the main modes of the flexure mechanisms, which are below 100 Hz. The system was controlled using a DSpace 1103 real-time control system with a loop-rate of 1 kHz. As the sampling rate is the same as the loop-rate of the control system, this setting provided a Nyquist frequency of 500 Hz, which is sufficient for the estimation of the frequency response below 300 Hz (including the un-modeled modes shown as Table 5-1).

5.2.2. Experimental Results and Discussion

To conduct the experiment, the frequency response from the actuator force P to the X direction displacement (X_1) of the motion stage (see Fig 5-1) was examined. The system identification was conducted in an open loop architecture. An up-chirp sinusoid signal from 0.1 Hz to 500 Hz was used as the command signal and was sent to the actuator. The reason to use chirp signal instead of white noise signal is that the former approach has lower uncertainty error in the estimation of the frequency response, since chirp signal is more concentrated in the frequency range of interest. The time domain response of X-axis displacement was collected. The frequency response was then obtained by comparing the frequency spectrum of the input and output signals, using MATLAB spectrum analysis with frequency-dependent resolution (spafdr).

A comparison between model prediction and experimental result is presented in Fig 5-9, using a representative operating condition where mass asymmetry is +10% and operating point is 8% of L (beam length, 47.5 mm). The damping ratios of the model were estimated based on experimental results, using the model of Rayleigh damping [57]. The estimated damping ratios were 0.009 for the “rigid-body” mode and 0.004 for the closely spaced modes (modes of interest). The “rigid-body” mode has a larger damping ratio since it is associated with the oscillation of the

entire flexure mechanism, which leads to a larger amount of area exposed to the viscous air environment. Note that the back electromagnetic force (back EMF) of the voice coil motor does not contribute to the damping ratio of the "rigid-body" mode when the current loop is turned on.

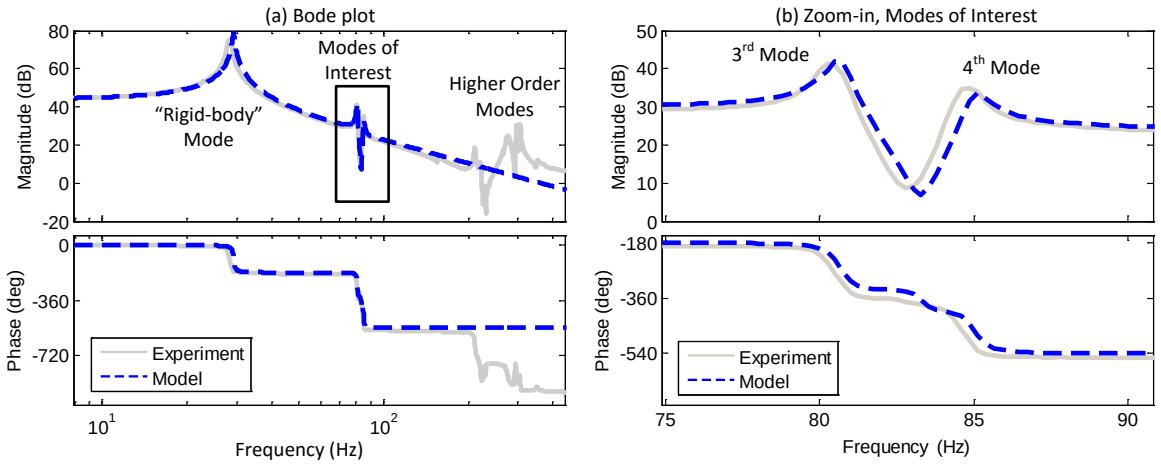

Fig 5-9 Experimental frequency response, asymmetry = 10%, operating point = 8%: (a) Bode plot; (b) zoom-in Bode plot of the modes of interest.

Key observations on Fig 5-9 are presented as below. First, the first four modes in the experimental results match well with the theoretical model (SR4DoF). Furthermore, the change in the frequencies of the 3rd and 4th modes due to variation of the operating point mostly agrees with the model prediction (not shown in Fig 5-9). Second, the modes of interest (the 3rd and 4th modes) are shown to be well-separated from other modes, as expected from the design stage. A factor of 2.5 in frequency separation was achieved in both directions. Moreover, it can be seen from Fig 5-9 that, the higher order modes (i.e., un-modeled modes) have a minor contribution on the dynamics at low frequency region. Third, CMP-CNMP zeros were observed as predicted by the theoretical model (SR4DoF). It can be seen from Fig 5-9b that the 360° phase drop due to the existence of CMP-CNMP zeros are observed.

In Fig 5-10, multiple experimental results were overlaid on the model predictions (the CNMP map in Fig 5-2). The two variables for parameter variations are the operating point and the amount of intended asymmetry. The two-variable testing grid has a size of 15 by 10 points. Note that only half of the area is tested due to the symmetric properties of the flexure w.r.t. positive/negative operating points. The squares indicate the conditions occurring when CMP-

CNMP zeros are present in the frequency response. As contrast, the solid circles indicate conditions occurring when CMP-CNMP zeros are absent (i.e., zeros are minimum phase). Lastly, the dashed circle represents the parameter set presented in Fig 5-9.

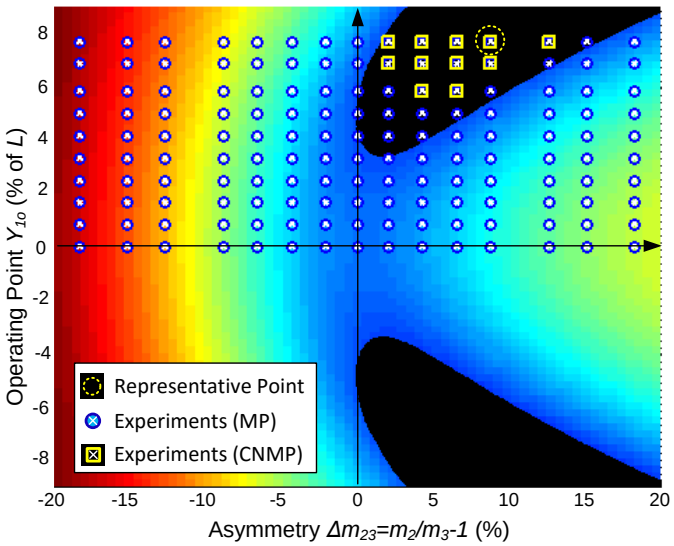

Fig 5-10 Simulation and experimental results of CNMP map.

Fig 5-10 shows that the experimental results match well with the model prediction as described below. First, for the situation of negative asymmetry ($\Delta m_{23}<0$), the entire operating range is free of CMP-CNMP zeros. Second, for $\Delta m_{23}>0$, experimental findings largely agree with model prediction of the CMP-CNMP zero region and confirm that CMP-CNMP zeros appear only when the operating point is larger than a certain value. Third, when the situation is nominally symmetric ($\Delta m_{23}\approx 0$), the frequency response is very sensitive to mass parameter variation in terms of the existence of CMP-CNMP zeros at large operating points.

The minor mismatch between the experimental and predicted CMP-CNMP region in Fig 5-10 could be attributed to the followings.

1. The CNMP map compiled from the theoretical model was based on the assumption of neglecting the damping ratios (see Assumption 3-2). However, as shown in chapter 3.3.1, the presence of non-zero damping changes the zero loci, with a tendency of moving all of the loci to the left side of the complex plane. This influence is most apparent when the zeros of interest are closed to the imaginary axis. In other words, the presence of non-

zero damping "delays" the zeros to transition from MMP zeros to CMP-CNMP zeros. This conclusion agrees with the experimental results shown in Fig 5-10. In Fig 5-10, when both the asymmetry and the operating point are small (e.g., asymmetry = 2%, operating point = 4%), the predicted CMP-CNMP zeros are not observed in experiments.

2. There are inevitable manufacturing tolerances that lead to an imperfect symmetry between stages ② and ③. A comparison between the theoretical model and the experimental result is shown in Fig 5-11. The operating conditions are asymmetry = 0% and operating point = 8%. In theoretical model, since asymmetry is zero, the 4th mode is completely unobservable regardless of the operating point. However, both the 4th mode was observed in the experiment, indicating a non-zero mass difference between stages ② and ③. The estimated initial mass difference is -2%, which indicates that stage ③ is slightly heavier than stage ②. This initial mass difference could shift the overall measurements (Fig 5-10) horizontally.

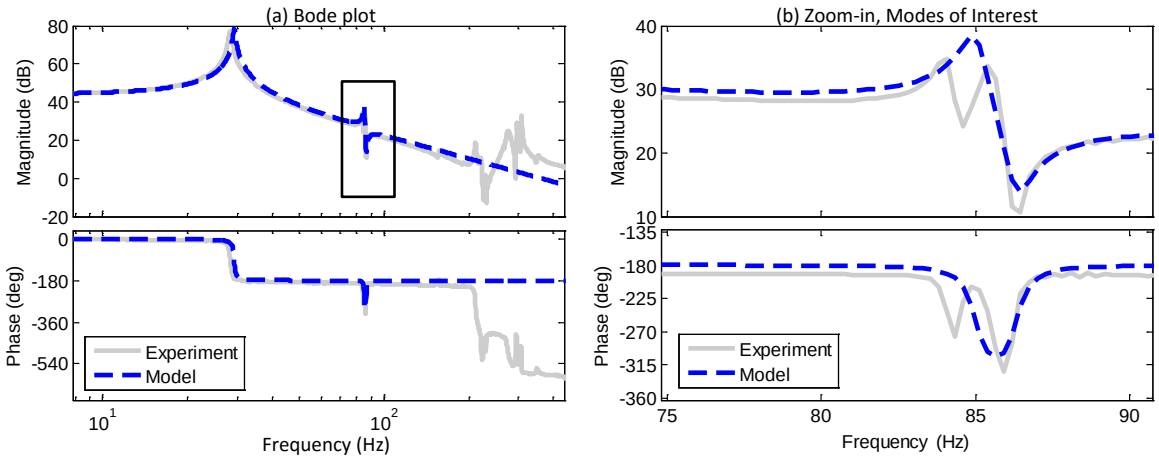

Fig 5-11 Experimental frequency response, asymmetry = 0%, operating point = 8%: (a) Bode plot; (b) zoom-in Bode plot of the modes of interest.

3. Other factors that could affect the measurements are: the parameter grid resolution used in the testing, and the scaling error due to the accuracy of visual measurement of the angle used in the virtual pulley.

5.3. Conclusion

The conclusions are summarized as below.

1. This work presents the first experimental validation of a mathematical model that predicts CNMP zeros under certain parameter space. Via this validation, we are confirmed that it is the geometric arc length conservation that leads to the coupling between motions in different axes. And this coupling, together with structural asymmetry, leads to interaction of the closely spaced modes. Such interactions under a non-zero contribution of the modes other than the modes of interest can give rise to CNMP zeros. This validation supports the modeling work and indirectly justifies the assumptions made in the modeling process (see chapter 3 and 4).
2. This work presents several careful design considerations in the setup and the experimental procedure to enable the above validation. Via these considerations, we successfully fabricated an experimental setup whose major dynamic behavior is the same as the theoretical model. The un-modeled modes (e.g., the out-of-plane modes and the rotational in-plane modes) were effectively arranged to high-frequency region that is irrelevant to the interest of this validation. Therefore, the development procedure from the theoretical model to experimental setup can be potentially useful in flexure mechanism design.
3. There always exists the tradeoff between making the stage more rigid and reducing the weight of the stage without expensive approaches (e.g., using truss-like structure). The key to balance this tradeoff is to come up with a reasonable specification. In this study, reducing the size of the secondary stage in the minor DPFM could help increase the associated natural frequencies. But this also leads to lower axial stiffness and creates additional compliance mode. However, as long as latter modes are still in irrelevant high frequencies region, increasing the frequencies of the former ones is more beneficial. However, in another example mainly involves quasi-static properties [3, 100] , making the stage more rigid become the key consideration.

Chapter 6. Mechatronic Design Considerations

So far, we have learned that the presence of CNMP zeros indicates that there are two closely spaced modes near the frequency of the zeros (in this thesis, CNMP zeros are mostly referred as CMP-CNMP quartet). These two modes are out-of-phase and lead to two resonance circles in the Nyquist plot with opposite directions. These circles are potentially problematic for closed loop robustness, if one wants to increase the close loop bandwidth.

In this chapter, two design strategies in setting the closed loop bandwidth are first compared with simulations, based on a system with two closely spaced modes (out-of-phase) and a double integrator as the remainder (Section 3.2.4). These results are then verified via the experimental setup developed in Chapter 5. It is shown that a system with a sequence of alternating signs is easier to stabilize than a system with non-alternating signs. Moreover, by intentionally setting up an asymmetric mass configuration, one can guarantee that the system has alternating signs. Consequently, a higher bandwidth is achieved by phase stabilization, compared to the cases with other mass configurations.

6.1. Two Strategies of Setting the Bandwidth

6.1.1. Bandwidth Lower than the First Resonance Frequency (Slow Control)

A simplified dynamic characteristic of a flexure mechanism is a second order mode (first resonance) plus other higher order modes (assumed well-separated from the first mode). Regarding the first resonance, there are two main strategies of closed loop classical control methods on setting the bandwidth: a bandwidth lower than the first resonance frequency (denoted as slow control), or a bandwidth higher than the first resonance frequency (denoted as fast control). The former strategy is considered in this section, which is applied in the field of

MEMS devices [101] and flexure mechanisms such as optical systems in satellites where damping is almost zero [78].

Two main advantages of the first strategy are as follows. First, high order modes of the plant can be effectively suppressed, thus good robustness is obtained against uncertainties. Second, by creating a favorable phase (usually with a low pass filter), the first resonance of the plant can be "actively damped" [59, 81]. An example is shown below. Consider the plant as Eq. (6-1) with a feedback control loop as in Fig 6-1.

$$P(s)=\frac{\omega_r^2}{s^2+2\zeta\omega_r s+\omega_r^2}=\frac{(10)^2}{s^2+2\cdot 0.01\cdot 10\cdot s+(10)^2} \tag{6-1}$$

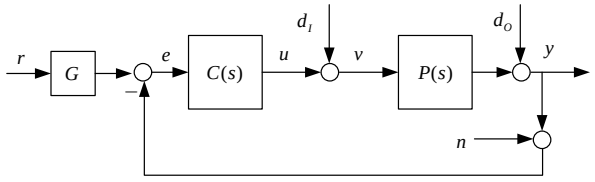

Fig 6-1 Classic control loop with unity feedback.

Consider a controller as Eq. (6-2).

$$C(s)=\frac{K}{s}\cdot\frac{\omega_{lp}^2}{s^2+2\zeta_{lp}\omega_{lp}s+\omega_{lp}^2}=\frac{1}{s}\cdot\frac{(5)^2}{s^2+2\cdot 0.7\cdot 5\cdot s+(5)^2} \tag{6-2}$$

The controller is an integrator with a 2nd order low pass filter (one can also consider replacing the single integrator with a double integrator). The low pass filter is set at half of the frequency of the first resonance with a damping ratio of 0.7. The open loop transfer function *L* is shown in Fig 6-2a. The crossover frequency is then set with a factor of 0.1 of the plant's first resonance frequency.

To illustrate the idea of "actively damping", the Nyquist plot of *L* is shown in Fig 6-2b. Consider different damping ratios of the plant *P(s)*. One can see that the modulus margin is unchanged regardless of the damping ratio. This design is useful for very lightly damped systems. Note that the term "active damping" in this example is not increasing the damping ratio itself. Instead, it suppresses the magnitude of the 1st resonance peak to be around 0 dB in the closed

loop transfer function *T*, as shown in Fig 6-2c. Moreover, the smaller the damping ratio, the closer the closed loop gain is to 0 dB.

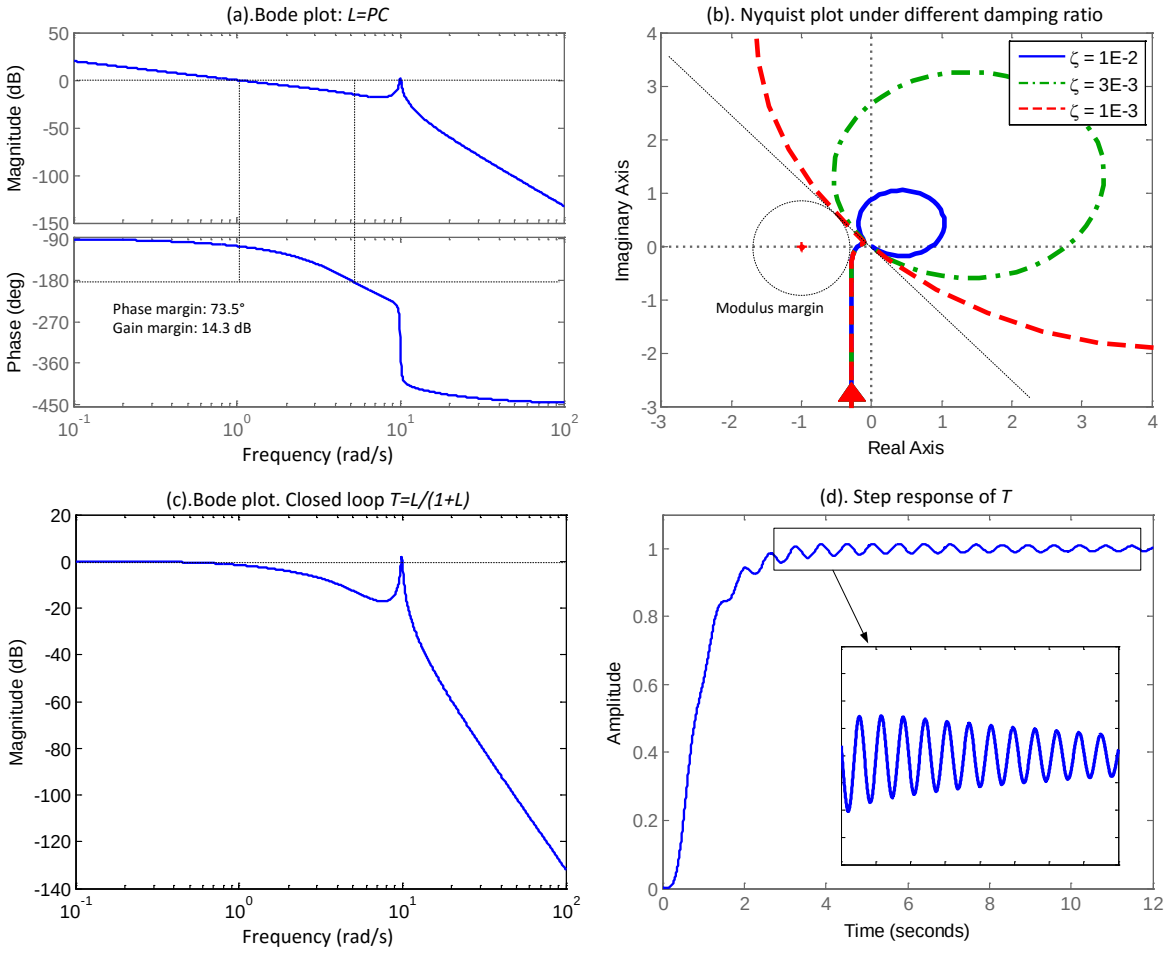

Fig 6-2 Characteristics of controller with bandwidth lower than 1st resonance frequency: (a) Bode plot of *L*; (b) Nyquist plot of *L*, under different damping ratios; (c) Bode plot of *T*; (d) step response of T.

However, the drawback is obvious. First, the bandwidth is limited by the first resonance frequency. For the example above, the maximum bandwidth can only be achieved at 10% of the first resonance frequency. To further increase the bandwidth, one can mechanically increase the frequency of the first resonance, such as increasing the stiffness or reducing the mass corresponding to the "rigid-body" mode of the flexure mechanism. However, there are other trade-offs that need to be considered. For example, in order to double the first resonance frequency, one needs the stiffness to be four times the original stiffness. As the main stiffness is

increased, there is a significant increase in the force required for the actuator to maintain the same travel range and to increase the motion speed. Second, there is a residual vibration due to the 1st resonance frequency in the step response of *T* (Fig 6-2d). This residual vibration is lightly damped and could adversely impact the positioning ability (e.g., standard deviation).

6.1.2. Bandwidth Higher than the First Resonance Frequency (Fast Control)

In this section, the second strategy is considered. However, this strategy could only be achieved with two main consequences. First, there is a limitation on using integrators in a controller. To explain, the phase of the plant after the first resonance is already -180°. Thus, to avoid adding more phase lag, these integrators need to be compensated via zeros before the crossover frequency. However, the poles of the integrators will be trapped by these zeros, leading to low frequency poles in the closed loop transfer function. One can increase the open loop gain to move these poles closer to the zeros, thus reducing the magnitude of these low frequency oscillations to an acceptable level. However, a large open loop gain will amplify the plant's higher order dynamics as well as noise.

Another consequence, compared to the first class, is the requirement of a better understanding of the plant (i.e., a model that captures the first few modes precisely). Ideally, if the frequency region between the first mode and the next major mode is "clean", then the bandwidth is limited by the next major mode. However, due to the superposition of relevant modes, the presence of lightly damped CMP zeros or NMP zeros imposes complexities on the control problem. In this work, we consider a situation that has two closely spaced modes appearing together as the second major modes. This situation can be approximated as two modes with a double integrator, namely mode u, mode v, and mode R (Section 3.2.4). Therefore, the plant considered is shown as Eq. (6-3). For simplicity, we assume that all the modes have identical damping ratios, and these damping ratios are sufficiently small. For illustration, consider a numerical example where ω_u=10, ω_v=12, and ζ=0.001.

$$G(s) = \frac{\alpha_R}{s^2} + \frac{\alpha_u}{s^2 + 2\zeta\omega_u s + \omega_u^2} + \frac{\alpha_v}{s^2 + 2\zeta\omega_v s + \omega_v^2};\ \omega_v > \omega_u > 0;\ \alpha_R > 0 \tag{6-3}$$

As discussed in Chapter 3, the sequence of these modes' signs is critical to closed loop robustness, if the crossover frequency increases to approach the frequencies of the two closely spaced modes. Among the four sequences of signs, there are two situations where the two closely spaced modes are out-of-phase, including the situation that leads to CMP-CNMP zeros (or dual RMP-RNMP zeros). These two situations are considered in this section (i.e., $\alpha_R>0$, $\alpha_u<0$, $\alpha_v>0$, and $\alpha_R>0$, $\alpha_u>0$, $\alpha_v<0$).

6.1.2.1. Alternating Signs ($\alpha_R>0$, $\alpha_u<0$, $\alpha_v>0$)

Consider the numerical set where α_R=7.98, α_u=-1.65, α_v=1.52 in Eq. (6-3). The presence of CMP-CNMP zeros are illustrated in Fig 6-3a and Fig 6-3b. The dual circles are in opposite directions as illustrated in Fig 6-3c, with the slower mode pointing upwards (negative direction).

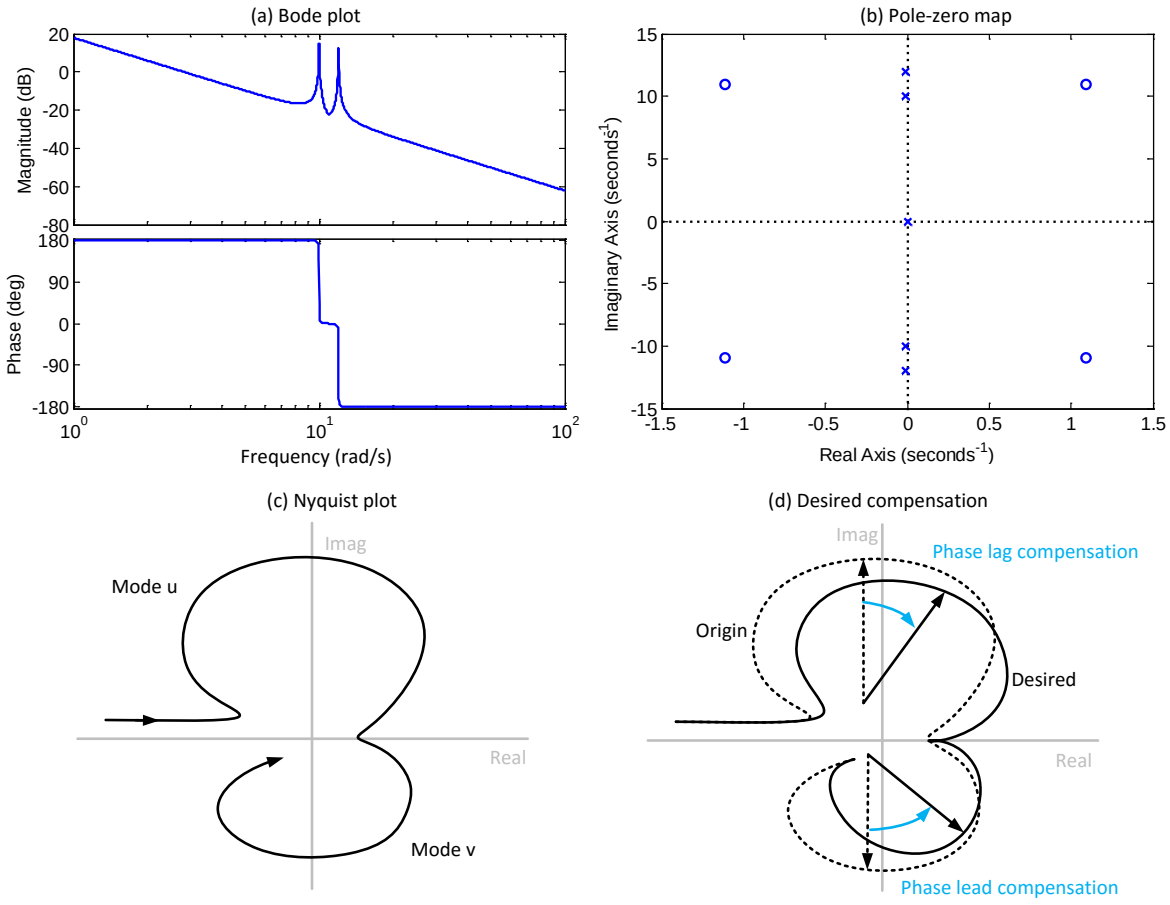

Fig 6-3 Properties of the numerical set α_R=7.98, α_u=-1.65, α_v=1.52: (a) Bode plot; (b) pole-zero map; (c) Nyquist plot; (d) desired compensation.

Based on the directions of the circles shown in Fig 6-3c, it is desirable to "rotate" the slower mode clockwise such that the corresponding circle grows away from the critical point (-1, 0). Similarly, it is desirable to "rotate" the faster mode counter-clockwise. Therefore, a phase lag compensation is desirable for the slower mode (mode u) while a phase lead compensation is desirable for the faster mode (mode v), as shown in Fig 6-3d.

A notch filter (Fig 6-4a) provides such desirable phase compensations as described above. Moreover, the notch filter provides an effective suppression to reduce the magnitude of the peaks. However, it is noteworthy that the notch filter has to be set such that mode u and mode v are located on opposite sides of the notch center.

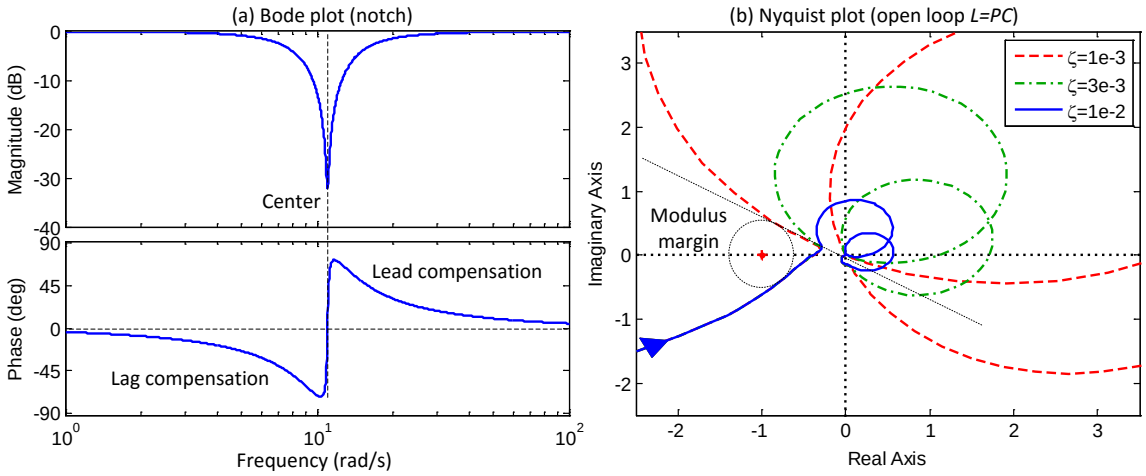

Fig 6-4 Compensator and open loop transfer function: (a) notch filter; (b) Nyquist plot of open loop *L*.

The controller considered for the plant in Fig 6-3a is shown as Eq. (6-4), which has the above-mentioned notch filter and a lead filter to create enough phase margin around the crossover frequency.

$$C(s)=\frac{1+s/1.82}{1+s/13.7}\cdot\frac{s^2+2s\cdot(0.01)\cdot 11+11^2}{s^2+2s\cdot(0.4)\cdot 11+11^2} \tag{6-4}$$

The Nyquist plot of the open loop transfer function *L=PC* is shown in Fig 6-3b, under different damping ratios. It can be seen that the modulus margin is unchanged since the circles are "rotated" to a safe angular position. In this example, the crossover frequency is around 4

rad/s (i.e., 40% of the frequency of the slower mode u), while the modulus margin is around 0.5 (i.e., the peak of the corresponding sensitivity function is bounded by 6 dB). More modulus margin can be obtained with a consequence of reducing the crossover frequency. One can see that there is some similarity between this approach and the one discussed in previous section. Moreover, the approach discussed in this example can be also applied to systems with the same sequence of signs but different types of zeros.

6.1.2.2. Non-alternating Signs (α_R>0, α_u>0, α_v<0)

Consider the numerical set where α_R=7.98, α_u=1.65, α_v=-1.52 in Eq. (6-3). Note that only the signs of α_u and α_v are changed compared to the previous section. The properties of this plant is shown in Fig 6-5a to Fig 6-5c. Note that the slower mode (mode u) is in the positive direction and thus points downwards in Nyquist plot (Fig 6-5c).

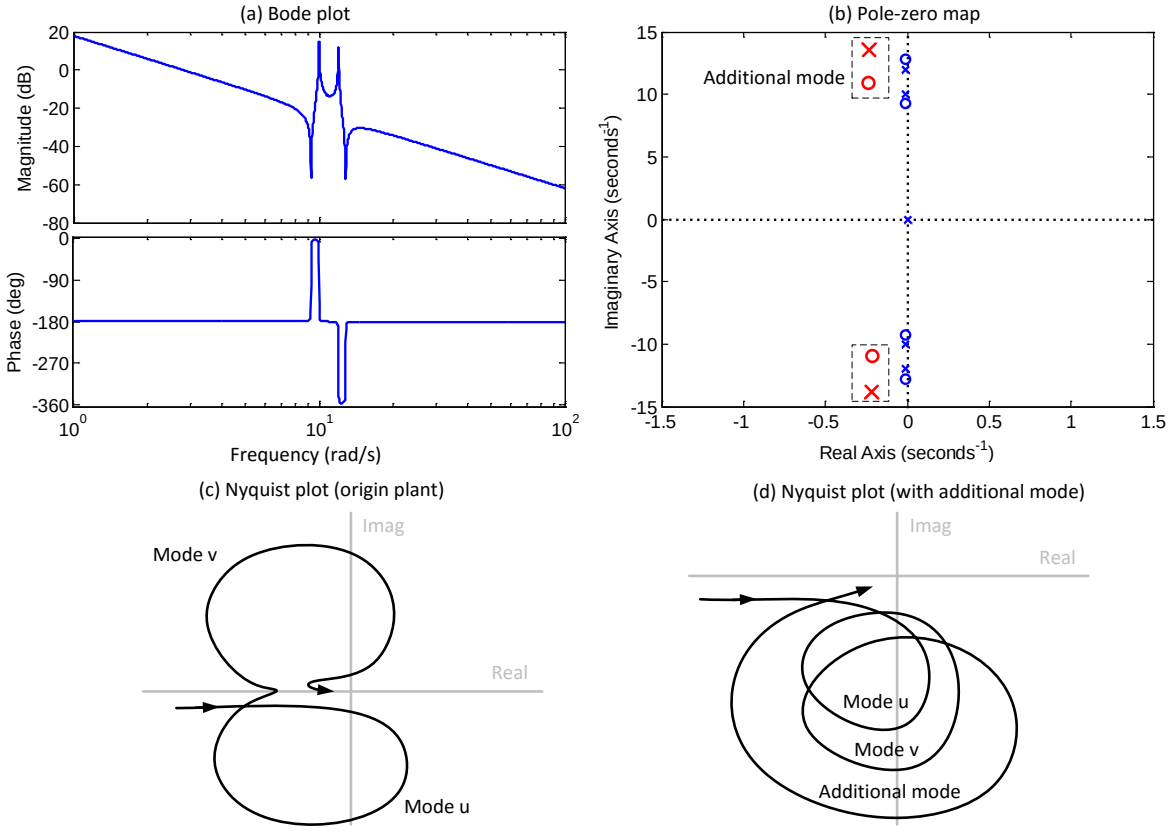

Fig 6-5 Properties of the numerical set α_R=7.98, α_u=1.65, α_v=-1.52: (a) Bode plot; (b) pole-zero map; (c) Nyquist plot; (d) Compensation.

The approach considered is attempted to achieve "phase stabilization" by adding an additional complex minimum phase pole-zero pair to the plant (shown as Fig 6-5b), such that both mode u and mode v are in the same directions (shown as Fig 6-5d). If a proper amount of phase lead compensation is applied to all of the three modes (including the addition mode), then one can "rotate" all of these three modes counter-clockwise to a safe angular position.

Consider the controller in Eq. (6-5), which has a lead filter and an additional 2nd order mode with a damping ratio of ζ. A comparison between two choices of ζ is shown in Fig 6-6. Note that all of the poles and zeros of the original plant have almost the same damping ratio as 0.001.

$$C(s) = \frac{1+s/1.82}{1+s/13.7} \cdot \frac{13^2}{11^2} \cdot \frac{s^2+2s\cdot\zeta\cdot 11+11^2}{s^2+2s\cdot\zeta\cdot 13+13^2} \tag{6-5}$$

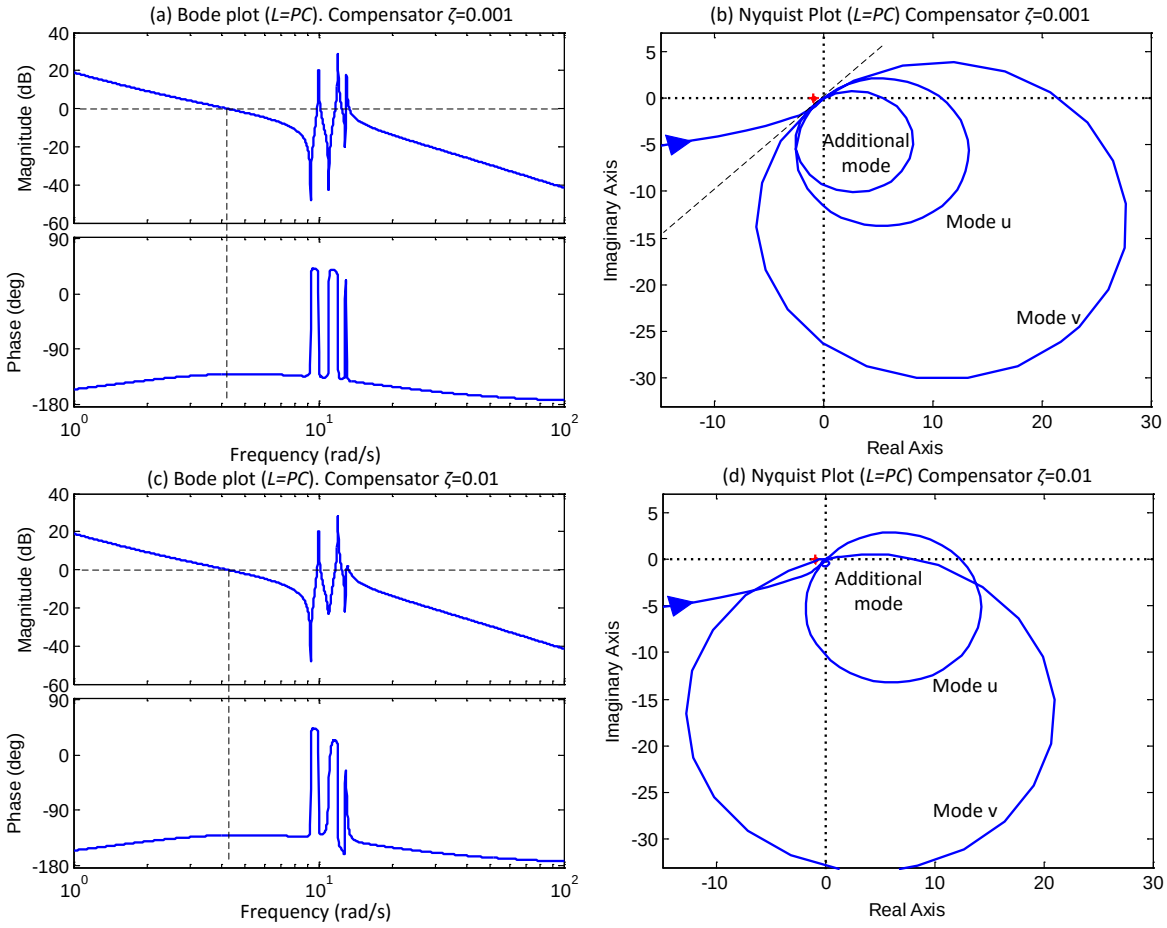

Fig 6-6 A comparison between controllers with different damping ratio: (a) Bode plot (ζ=0.001); (b) Nyquist plot (ζ=0.001); (c) Bode plot (ζ=0.01); (d) Nyquist plot (ζ=0.01).

From Fig 6-6a and Fig 6-6c one can see that the lead filter provides approximately a constant phase compensation around 10 rad/s (about 45° from -180°). In this situation, If the damping ratio of the additional mode is the same as those of the plant's poles and zeros, then all of the resonances start and end at the same phase (i.e., -135°, see Fig 6-6a). As a result, all of the circles start and end at the same angular position, growing away from the critical point (-1, 0), as shown in Fig 6-6b. However, if the damping ratio of the additional mode is larger than those of the plant's poles and zeros, then there is not enough phase compensation for the faster mode (mode v), as shown in Fig 6-6c and Fig 6-6d. As a result, the circle corresponding to mode v cannot be "rotated" to a safe angular position, leading to a stability problem.

Therefore, if the phase compensation is constant around the frequencies of mode u and mode v, then it is important to have the damping ratio of the additional mode no larger than the damping ratio of mode u and mode v (i.e., the damping ratio of the plant's poles and zeros). A larger damping ratio of the additional mode may lead to an unsafe angular position of the circle, corresponding to the faster mode v. Moreover, if the phase compensation drops monotonously around the frequencies of mode u and mode v, then an even smaller damping ratio of the additional mode is required.

Based on the above-discussion, since the damping ratio of the closely spaced mode is very small in lightly damped systems (e.g., 0.004 in the experiment in Chapter5), it is difficult to both implement an additional mode with small damping ratio and precisely predict the damping ratios of the corresponding modes. Furthermore, the controller Eq. (6-5) also amplifies the magnitude in high frequencies due to the additional mode. Therefore, it is not recommended to consider the controller Eq. (6-5). Instead, it is recommended to apply a conventional controller (e.g., lead + 2^{nd} order low pass with 0.7 damping ratio), with the consequence of achieving a lower bandwidth (usually 10% of the slower mode u). However, the conventional approach requires a larger "clean" frequency region before the two closely spaced modes. In Section 6.2, we will see that it is difficult to implement a conventional controller since the "clean" region is limited.

6.2. Implementation based on SR4DOF

6.2.1. Model-based Controller Design

Consider the SR4DoF model discussed in Chapter 2 (shown in Fig 2-2), but with the dimensions used in Chapter 5. Fig 6-7 demonstrates a comparison between the model and experimental results, using mass asymmetry as +10% and the operating point as 8% of L (beam length, 47.5 mm). The damping ratios of the model are adjusted based on experimental results by applying a damping matrix proportional to the stiffness matrix (i.e., Rayleigh damping). CMP-CNMP zeros are observed. The first resonance frequency is measured at 179 rad/s with a damping ratio of 0.009; the closely spaced modes (i.e., mode u and mode v) are measured with almost the same damping ratios as 0.004. One can see from Fig 6-7a that there are high frequency modes with significant peaks after 1000 rad/s. These high frequency modes have to be suppressed effectively for robustness consideration.

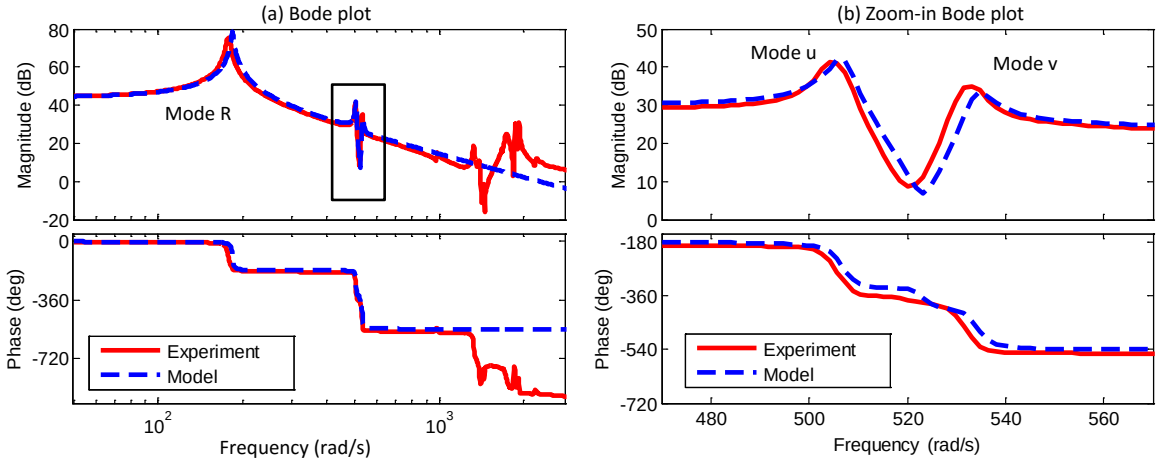

Fig 6-7 Frequency response of SR4DoF, model vs. experiment: (a) Bode plot; (b) zoom-in version.

The first controller tested is stated as Eq. (6-6) based on Section 6.1.1.

$$C(s)=\frac{K}{s}\cdot\frac{\omega_{lp}^2}{s^2+2\zeta_{lp}\omega_{lp}s+\omega_{lp}^2}=\frac{0.12}{s}\cdot\frac{(179/2)^2}{s^2+0.7\cdot 179\cdot s+(179/2)^2} \tag{6-6}$$

Based on the model discussed above, the Bode plot and the Nyquist plot of the open loop transfer functions are shown in Fig 6-8a and Fig 6-8b, respectively. One can see that the bandwidth is 10% of the first resonance (i.e., 18 rad/s), with a phase margin of 70°.

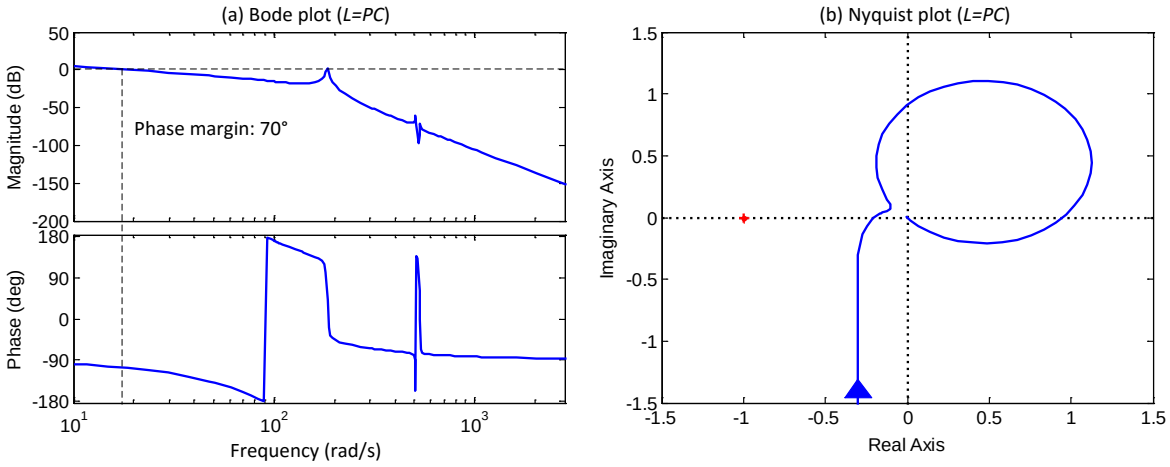

Fig 6-8 First controller (slow) open loop transfer function: (a) Bode plot; (b) Nyquist plot.

To increase bandwidth, the sequence of the decomposed modes' signs is considered first. Based on Chapter 2 and Chapter 5, CMP-CNMP zeros are observed when asymmetry is positive, while CMP-CNMP zeros are eliminated when asymmetry is negative. These are illustrated in Fig 6-9. Based on Chapter 3, one can see that alternating signs (i.e., $\alpha_R>0$, $\alpha_u<0$, and $\alpha_v>0$) are obtained when asymmetry is positive, while non-alternating signs (i.e., $\alpha_R>0$, $\alpha_u>0$, and $\alpha_v<0$) are obtained when asymmetry is negative.

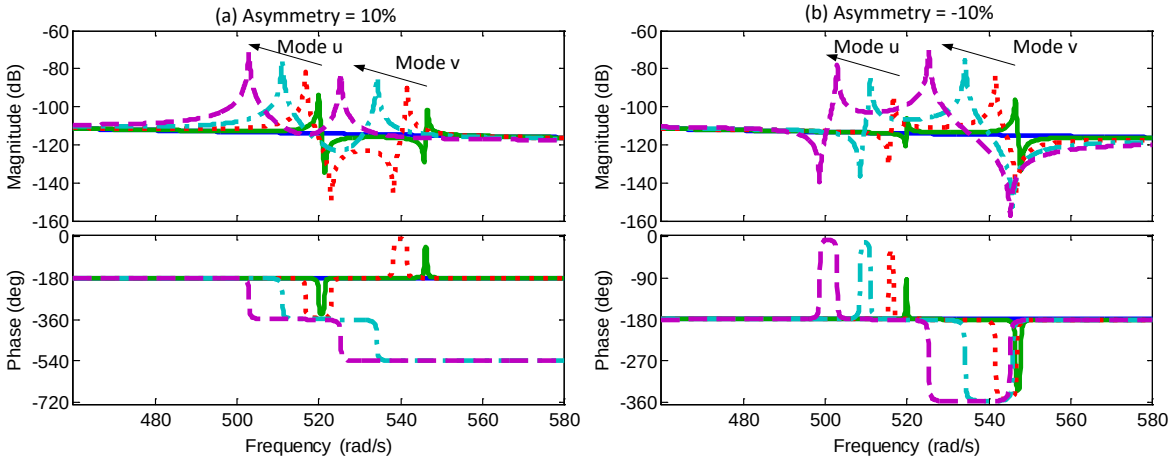

Fig 6-9 The closely spaced modes under an increasing operating point: (a) asym>0; (b) asym<0.

Based on Section 6.1.2, since the closely spaced modes are measured around 500 rad/s and the 1st mode is measured at 178 rad/s, there is a factor of 2.5 as frequency separation. Therefore, the situation of alternating signs (i.e., $\alpha_R>0$, $\alpha_u<0$, and $\alpha_v>0$) is favorable. Thus asymmetry is chosen with 10% for illustration. The controller tested is stated as Eq. (6-7).

$$C(s)=\frac{51110}{s}\cdot\frac{s+88.7}{s+625}\cdot\frac{1}{s+893}\cdot\frac{s^2+2s\cdot 0.4\cdot 175+175^2}{s^2+2s\cdot 0.7\cdot 700+700^2}\cdot\frac{s^2+2s\cdot 0.01\cdot 520+520^2}{s^2+2s\cdot 0.1\cdot 555+555^2} \tag{6-7}$$

In Eq. (6-7), there is a lead (88.7 rad/s and 625 rad/s); a 1st order low pass filter (893 rad/s); a complex zero-pair (175, damping 0.4) to provide additional phase margin; a 2nd order low pass filter (700 rad/s, damping 0.7); and a notch (520 rad/s, the poles are shifted to 555 rad/s in order to provide additional phase margin for crossover). The Bode plot and the Nyquist plot of the open loop transfer function are shown in Fig 6-10a and Fig 6-10b, respectively. The crossover frequency is set at 287 rad/s with a phase margin of 42°.

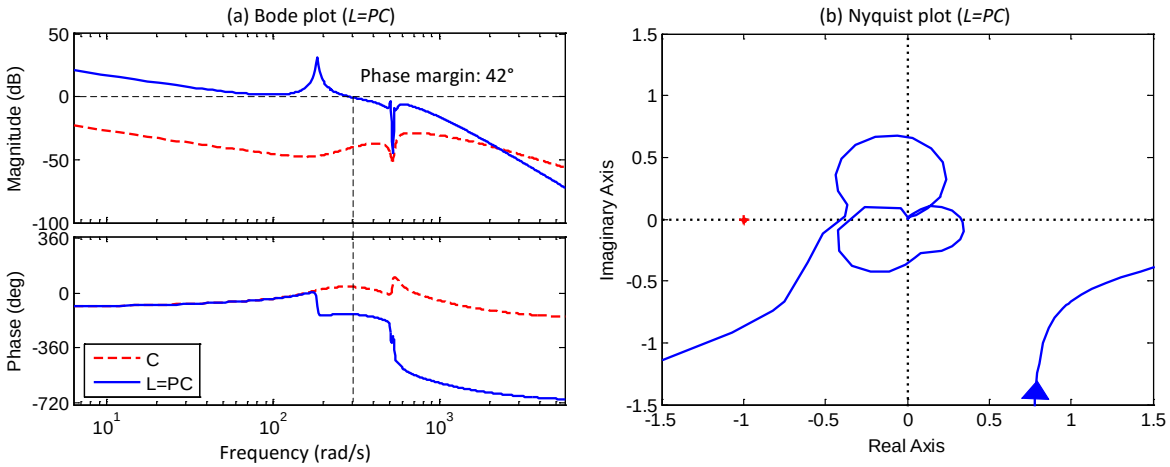

Fig 6-10 First controller (fast) open loop transfer function: (a) Bode plot; (b) Nyquist plot.

A comparison between the slow controller and the fast controller is provided in Fig 6-6. From the closed loop transfer function shown in Fig 6-6a, one can see that the fast controller has a certain frequency region below 0 dB before the crossover frequency. This is due to the region closed to 0 dB before the 1st resonance frequency as shown in Fig 6-10a, which is a result of the three zeros smaller than the 1st resonance frequency in the controller shown in Eq. (6-7). This "deep" in the closed loop transfer function will slow down the settling time in the step response.

The step responses of the two controlled systems are shown in Fig 6-10b. The system with the slow controller has a settling time of 0.1663 seconds while the system with the fast controller has a settling time of 0.0858 seconds.

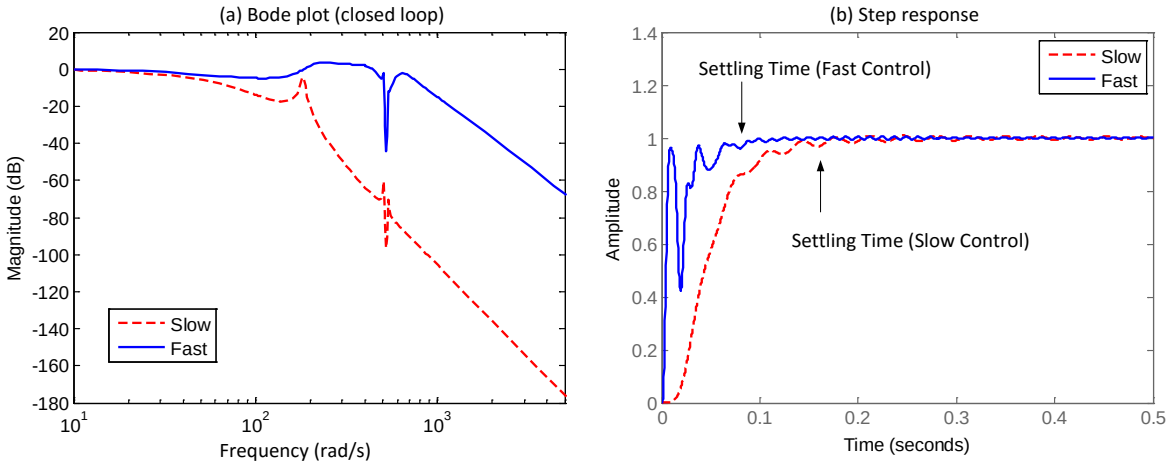

Fig 6-11 Comparison between fast and slower controller: (a) Bode plot; (b) step response.

6.2.2. Experimental Implementation

In this section, we apply the model-based controllers to the experimental setup developed in Chapter 5. A systematic comparison is shown in Fig 6-12. These plots are generated using the ideal controllers and the frequency response of the experimental setup. One can see a good match between the results from a physical model and the ones from a real plant, which again validates the model developed in Chapter 2. In the fast control case, the dual circles corresponding to CMP-CNMP zeros are “rotated” to safe angular positions as shown in Fig 6-12d. The third major circle in Fig 6-12d is due to the non-collocation between the poles and the zeros of the notch filter. Since the damping ratio of these poles is controllable (i.e., 0.4 in this example), this circle is not a major stability problem.

It is noteworthy that for the fast controller (Fig 6-12c), the high order modes have significant amplitudes, which corresponds to the various circles in mixed directions as shown in Fig 6-12d. These modes can lead to stability problem if one further increases the loop gain.

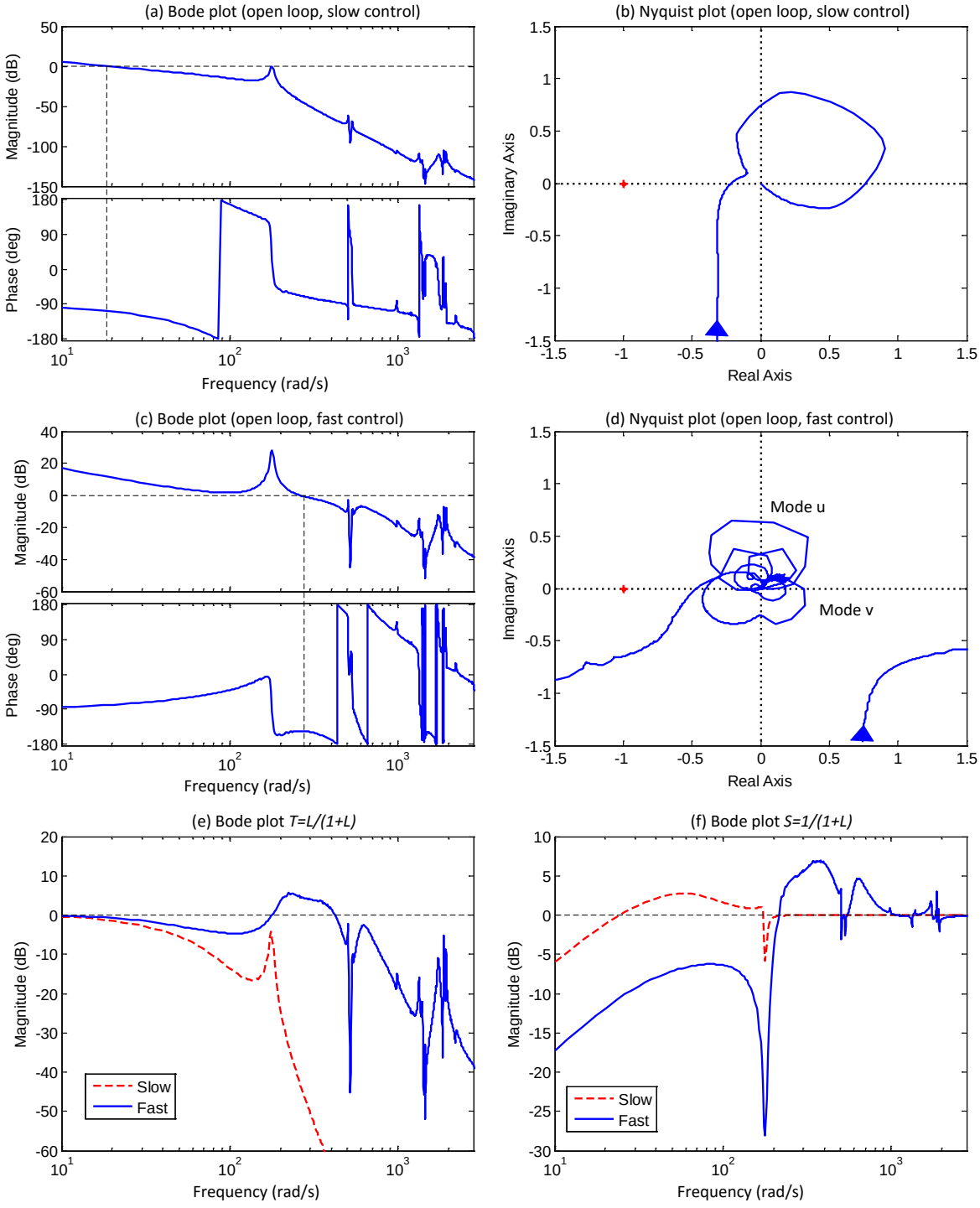

Fig 6-12 Comparison between fast and slow controller (estimated results with setup): (a) Bode plot (open loop, slow control); (b) Nyquist plot (open loop, slow control); (c) Bode plot (open loop, fast control); (d) Nyquist plot (open loop, fast control); (e) Bode plot (closed loop transfer function); (f) Bode plot (sensitivity function).

The step response of the two control methods are compared in Fig 6-13a, which agrees with the model prediction shown in Fig 6-10b. The distribution of the positioning error in a steady state is shown in Fig 6-13b. The data collected for the estimation of the probability density has 16,000 points (i.e., 3 seconds with a sampling time of 0.2 milliseconds). Thus, it is shown that the fast controller has smaller standard deviation (0.050 μm) than the slow controller (0.066 μm). This result is consistent with Fig 6-12f, indicating that the fast controller has better disturbance rejection ability at low frequency than the slow controller.

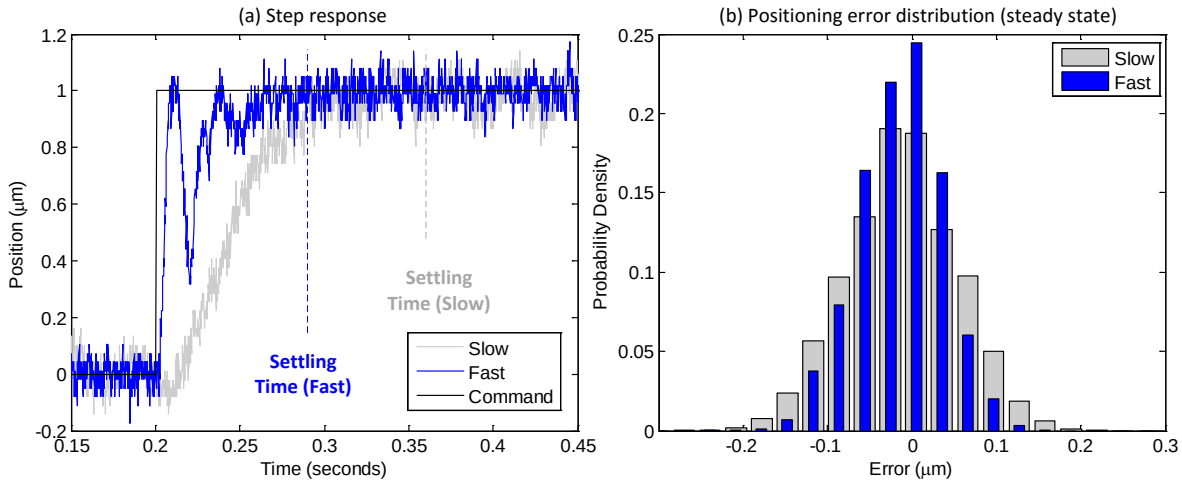

Fig 6-13 Experimental comparison between fast and slow control: (a) step response; (b) positioning error of steady state. Note that the bars indicate the quantization of the sensor, with a size of 0.030 μm.

Next, the fast controller is examined over the range of the operating point (Y_{1o}). In this test, the range is defined from 0 to 8% of L. We found that the standard deviation of the positioning has a minor change over the workspace, with the best achieved at Y_{1o}=0 as 0.048 μm and the worst achieved at Y_{1o}=8% as 0.050 μm. This is because the circle corresponding to mode u has smaller magnitude and a safer angular position at Y_{1o}=0. As the operating point increases, the magnitude increases and the circle "rotates" counter-clockwise, thus degrading the steady state standard deviation. However, since mode u and mode v have damping ratios of about 0.004, the size of the circle is not significant. Thus, such degradation is not apparent in experimental results. On the other hand, this observation highlights the importance and effectiveness of adding damping to the two closely spaced modes for robustness consideration.

6.3. Considerations on Mechanical System Design

Based on the above discussions and going back to the mechanical system design level, there are two main directions in improving the plant dynamics.

First, one can re-design the resonance frequencies. Rather than simply increasing the first resonance frequency, we can consider lowering the first resonance frequency while increasing the second major resonance frequency. Therefore, it is favorable to create a “clean” segment, similar to procedure discussed in Chapter 5 [102]. There are already some recent studies that look promising. For example, Panas shows a nested linkage that can effectively increase the resonance frequency of a mode associated with the under-constraint effect [103]. This idea is implemented in another XY-stage design [104]. However, the consequence is a larger setup size and extra undesired modes due to the additional mass of the nested linkage.

The second direction is to increase the damping ratio of the relevant lightly damped modes. Increasing the damping ratio of these lightly damped modes will result in smaller size of the corresponding resonance circles and increase the speed of decaying the corresponding residual oscillations. The increase in the overall damping ratio for all higher order flexible modes will also improve the robustness [78]. There is also some research on damping methods, such as adding a sandwich viscous layer [105] or using the eddy current [106].

Lastly, even if it is theoretically possible to increase the bandwidth, there are other mechatronics problems need to be solved. These problems include the following: (1) environment conditioning to control the noise and disturbance down to an acceptable level; (2) choosing or developing proper actuators, sensors, and amplifiers, and (3) other associated control algorithms like command shaping and feedforward control. The advancement of the overall performance requires the improvement in all associated areas. The shortage on any of these areas will become the next bottleneck. The discussion in this section (plant dynamics and feedback controllers) is just on one of these such areas.

6.4. Conclusion

In this chapter, two strategies of setting the closed loop bandwidth are considered. These two strategies are characterized by the frequency of their corresponding bandwidth w.r.t. the 1st resonance frequency, namely slow control and fast control. The slow control has good robustness against high frequency uncertainties while suffers from low bandwidth. The fast control has a higher bandwidth but requires a good understanding of the high frequency dynamics.

Regarding the fast control, two types of systems are considered. Both systems have the same system type (i.e. two closely spaced modes that are out-of-phase and a double integrator as a remainder) but different sequence of the decomposed modes' signs. It is shown that in lightly damped systems, the system type with a sequence of signs that are alternating has better robustness against uncertainties in damping ratios.

The SR4DoF model in Chapter 2 and the experimental setup in Chapter 5 are used for validation. It is shown that by assigning intentionally positive/negative mass asymmetry, one obtains a system with alternating/non-alternating signs, respectively. Therefore, positive asymmetry is favorable, even in the case when CMP-CNMP zeros are observed. This finding is counter-intuitive since NMP zeros are generally unfavorable. It is shown that it is the sequence of signs rather than the presence of NMP zeros that determines whether the plant can be controlled with better robustness. The presence of NMP zeros, instead, is the consequence of certain sequence of dominant modes and the sequence of the decomposed modes' signs (as discussed in Chapter 3).

Chapter 7. Conclusion and Future Work

7.1. Conclusion

In, Chapter 2, a lumped-parameter modeling approach is proposed to analytically model the dynamics of flexure mechanisms comprising the DPFM. This model captures the key relevant geometric non-linearity (i.e. arc length conservation) in large displacement flexure mechanics. Linearization about any arbitrary operating point enables frequency domain analysis. Based on this model, we are able to predict previously unexplained CNMP zeros seen experimentally. The model establishes the existence of CNMP zeros under certain combinations of operating point and parametric asymmetry in the non-collocated transfer function of a simple representative XY flexure mechanism. This finding helps generate the design insight that, rather than an intuitively symmetric design.

In Chapter 3 and 4, a theoretical model is proposed to explain the conditions (when), the mechanisms (how), and the physical interpretations (why) on the occurrence of the CNMP zeros in flexure mechanisms. The zero crossing and undershoot of the step response of a system with real/complex NMP zeros are explained in view of modal decompositions. The root locus technique is used to find the migration of zeros. Particularly for CNMP zeros, the impact of damping ratio is discussed. As CNMP zeros and curve veering both exist in closely spaced modes, the connection between CNMP zeros and curve veering is explored.

In Chapter 5, the above modeling work is validated experimentally. The experiment is carefully designed to isolate the features of interest (i.e. the closely spaced modes) from other factors (e.g. parasitic dynamics). The existence of CNMP zeros are validated in lightly damped, passive stable flexure mechanisms. This work also shows that via an intentionally asymmetric structure, the CNMP zeros can be eliminated completely.

In Chapter 6, the knowledge developed in Chapter 2-4 is used in a mechatronics design consideration, with the purpose to increase closed loop bandwidth of flexure based motion stages.

Using the SR4DoF model as an example, it is shown that if the two closely spaced modes are out-of-phase, a system with a sequence of alternating signs is easier to stabilize than a system with a sequence of non-alternating signs, even when CNMP zeros appear in the former one. Additionally, an intentionally positive mass asymmetry configuration can guarantee a sequence of alternating signs. In both simulation and experiments, we demonstrate a closed loop control with the bandwidth around the 1st resonance frequency with positive mass configuration.

7.2. Future Work

Based on discussion in this thesis, three main future directions are considered. All these future tasks indicate the next significant breakthrough on the overall performance of flexure based motion stages.

First, for establishing physical models (i.e., Chapter 2), the SR4DoF model is a simplified model of the XY flexure stage. It is valuable to understand the behavior of a more complicated XY flexure stage. This requires a further study on a higher dimensional model.

Second, for understanding the zeros on a more fundamental level (i.e., Chapter 3), it is valuable to extend the investigation to a more general system (e.g., non-zero damping systems or continuous systems). In addition, it is of great interest to study the propagation of energy from the actuator to the sensor, which paves the road to further understand the physics of NMP zeros.

Lastly, as shown in Chapter 6, other than simply looking for a high first resonance frequency of a flexure system, it is desired to have a design that has lower first resonance frequency, higher second major resonance frequency, and “clean” dynamics in between the first two major modes. In addition, adding passive damping can help stabilize the system with a more aggressive controller. Further methodologies can be obtained to form a more detailed design guideline.

Bibliography

[1] L. C. Hale, "Principles and techniques for designing precision machines," Lawrence Livermore National Lab., CA (US)1999.

[2] S. T. Smith, *Flexures: elements of elastic mechanisms*. Amsterdam : Abingdon: Gordon & Breach ; Marston, 2000.

[3] S. Awtar, "Synthesis and Analysis of Parallel Kinematic XY Flexure Mechanisms," MIT, 2003.

[4] J. B. Hopkins, "Design of parallel flexure systems via freedom and constraint topologies (FACT)," Massachusetts Institute of Technology, 2007.

[5] B. P. Trease, Y.-M. Moon, and S. Kota, "Design of large-displacement compliant joints," *Journal of mechanical Design,* vol. 127, pp. 788-798, 2005.

[6] D. K. Blanding, *Exact Constraint: Machine Design Using Kinematic Principles*. New York, NY: ASME Press, 1999.

[7] T. R. Tauchert, *Energy principles in structural mechanics*: McGraw-Hill NY, 1974.

[8] Y. Yong, S. Moheimani, B. J. Kenton, and K. Leang, "Invited Review Article: High-speed Flexure-guided Nanopositioning: Mechanical Design and Control Issues," *Review of Scientific Instruments,* vol. 83, p. 121101, 2012.

[9] H. Sadeghian, N. Koster, and T. van den Dool, "Introduction of a high throughput SPM for defect inspection and process control," in *SPIE Advanced Lithography*, 2013, pp. 868127-868127-8.

[10] A. Mohammadi, A. G. Fowler, Y. K. Yong, and S. R. Moheimani, "A feedback controlled MEMS nanopositioner for on-chip high-speed AFM," *Microelectromechanical Systems, Journal of,* vol. 23, pp. 610-619, 2014.

[11] S. Awtar and G. Parmar, "Design of a Large Range XY Nanopositioning System," *Journal of Mechanisms and Robotics,* vol. 5, pp. 021008-021008, 2013.

[12] G. Parmar, K. Barton, and S. Awtar, "Large dynamic range nanopositioning using iterative learning control," *Precision engineering,* vol. 38, pp. 48-56, 2014.

[13] G. K. Parmar, "Dynamics and Control of Flexure-based Large Range Nanopositioning Systems," 2014.

[14] S. William, "Control system: Advanced methods," ed: Taylor and Francis Group, LLC, 2011.

[15] J. S. Freudenberg and D. P. Looze, "Right half plane poles and zeros and design trade-offs in feedback systems," *IEEE Transactions on Automatic Control,* vol. 30, pp. 555-565, 1985.

[16] J. S. Freudenberg, *A first graduate course in feedback control*. University of Michigan, 2012.

[17] C.-C. Lan, K.-M. Lee, and J.-H. Liou, "Dynamics of highly elastic mechanisms using the generalized multiple shooting method: Simulations and experiments," *Mechanism and Machine Theory,* vol. 44, pp. 2164-2178, 2009.

[18] T. T. Akano and O. A. Fakinlede, "An Effective Finite Element Method for Beam Based Compliant Mechanism," *International Journal of Engineering and Technology,* vol. 3, pp. 407-419, 2013.

[19] V. Shilpiekandula and K. Youcef-Toumi, "Characterization of dynamic behavior of flexure-based mechanisms for precision angular alignment," in *2008 American Control Conference (ACC '08), 11-13 June 2008*, Piscataway, NJ, USA, 2008, pp. 3005-10.

[20] Y. Yue-Qing, L. L. Howell, C. Lusk, Y. Ying, and H. Mao-Gen, "Dynamic modeling of compliant mechanisms based on the pseudo-rigid-body model," *Transactions of the ASME. Journal of Mechanical Design,* vol. 127, pp. 760-5, 2005.

[21] C. Boyle, L. L. Howell, S. P. Magleby, and M. S. Evans, "Dynamic modeling of compliant constant-force compression mechanisms," *Mechanism and Machine Theory,* vol. 38, pp. 1469-1487, 2003.

[22] L. L. Howell, A. Midha, and T. Norton, "Evaluation of equivalent spring stiffness for use in a pseudo-rigid-body model of large-deflection compliant mechanisms," *Journal of mechanical Design,* vol. 118, pp. 126-131, 1996.

[23] S. K. Dwivedy and P. Eberhard, "Dynamic analysis of flexible manipulators, a literature review," *Mechanism and Machine Theory,* vol. 41, pp. 749-777, 2006.

[24] S. Timoshenko and J. Goodier, "Theory of elasticity, 1951," *New York,* vol. 412.

[25] E. Reissner, "On one-dimensional large-displacement finite-strain beam theory," *Studies in applied mathematics,* vol. 52, pp. 87-95, 1973.

[26] J. Jonker and J. Meijaard, "A geometrically non-linear formulation of a three-dimensional beam element for solving large deflection multibody system problems," *International Journal of Non-Linear Mechanics,* vol. 53, pp. 63-74, 2013.

[27] S. Sen, "Beam constraint model: Generalized nonlinear closed-form modeling of beam flexures for flexure mechanism design," University of Michigan, 2013.

[28] S. Awtar and S. Sen, "A Generalized Constraint Model for Two-Dimensional Beam Flexures: Nonlinear Strain Energy Formulation," *Journal of mechanical Design,* vol. 132, pp. 0810091-08100911, Aug 2010.

[29] L. L. Howell, *Compliant mechanisms*: John Wiley & Sons, 2001.

[30] S. Awtar, A. Slocum, and E. Sevincer, "Characteristics of Beam-Based Flexure Modules," *Journal of mechanical Design,* vol. 129, pp. 625-639, 2007.

[31] V. Modi, "Attitude dynamics of satellites with flexible appendages-a brief review," *Journal of Spacecraft and Rockets,* vol. 11, pp. 743-751, 1974.

[32] K. Pandalai, "Large Amplitude free flexural vibration of structures," *Journal of reinforced plastics and composites,* vol. 6, pp. 153-161, 1987.

[33] T. Kane, R. Ryan, and A. Banerjee, "Dynamics of a cantilever beam attached to a moving base," *Journal of Guidance, Control, and Dynamics,* vol. 10, pp. 139-151, 1987.

[34] J. Yang, L. Jiang, and D. C. Chen, "Dynamic modelling and control of a rotating Euler–Bernoulli beam," *Journal of Sound and Vibration,* vol. 274, pp. 863-875, 2004.

[35] M. R. M. C. Dasilva, "Non-Linear Flexural Flexural Torsional Extensional Dynamics of Beams .1. Formulation," *International journal of solids and structures,* vol. 24, pp. 1225-1234, 1988.

[36] B. Jonker, "A finite element dynamic analysis of spatial mechanisms with flexible links," *Computer Methods in Applied Mechanics and Engineering,* vol. 76, pp. 17-40, 1989.

[37] A. H. Nayfeh, "NONLINEAR TRANSVERSE VIBRATIONS OF BEAMS WITH PROPERTIES THAT VARY ALONG THE LENGTH," *Journal of the Acoustical Society of America,* vol. 53, pp. 766-770, 1973.

[38] L. D. Zavodney and A. H. Nayfeh, "The Non-Linear Response of a Slender Beam Carrying a Lumped Mass to a Principal Parametric-Excitation - Theory and Experiment," *International Journal of Non-Linear Mechanics,* vol. 24, pp. 105-125, 1989.

[39] H. Moeenfard and S. Awtar, "Modeling Geometric Nonlinearities in the Free Vibration of a Planar Beam Flexure With a Tip Mass," *Journal of mechanical Design,* vol. 136, p. 044502, 2014.

[40] R. S. Edmunds and D. Mingori, "Robust control system design techniques for large flexible space structures having non-colocated sensors and actuators," 1982.

[41] R. H. Cannon and D. E. Rosenthal, "Experiments in control of flexible structures with noncolocated sensors and actuators," *Journal of Guidance, Control, and Dynamics,* vol. 7, pp. 546-553, 1984.

[42] V. Spector and H. Flashner, "Sensitivity of structural models for noncollocated control systems," *Journal of Dynamic Systems, Measurement, and Control,* vol. 111, pp. 646-655, 1989.

[43] G. D. Martin, "On the control of flexible mechanical systems," Department of Electrical Engineering, Standford University, 1978.

[44] W. B. Gevarter, "Basic relations for control of flexible vehicles," *AIAA journal,* vol. 8, pp. 666-672, 1970.

[45] V. A. Spector and H. Flashner, "Modeling and design implications of noncollocated control in flexible systems," *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control,* vol. 112, pp. 186-93, 1990.

[46] D. K. Miu, "Physical interpretation of transfer function zeros for simple control systems with mechanical flexibilities," *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control,* vol. 113, pp. 419-24, 1991.

[47] R. H. Cannon and E. Schmitz, "Initial experiments on the end-point control of a flexible one-link robot," *The International Journal of Robotics Research,* vol. 3, pp. 62-75, 1984.

[48] M. Tohyama and R. H. Lyon, "Zeros of a Transfer-Function in a Multi-Degree-of-Freedom Vibrating System," *Journal of the Acoustical Society of America,* vol. 86, pp. 1854-1863, Nov 1989.

[49] N. Loix, J. Kozanek, and E. Foltete, "On the complex zeros of non-colocated systems," *Journal of Structural Control,* vol. 3, pp. 79-87, 1996.

[50] J. B. Hoagg, J. Chandrasekar, and D. S. Bernstein, "On the zeros, initial undershoot, and relative degree of collinear lumped-parameter structures," *Journal of Dynamic Systems, Measurement, and Control,* vol. 129, pp. 493-502, 2007.

[51] S. Awtar and K. C. Craig, "Electromagnetic coupling in a dc motor and tachometer assembly," *Journal of Dynamic Systems, Measurement, and Control,* vol. 126, pp. 684-691, 2004.

[52] R. S. Irving, *Integers, polynomials, and rings: a course in algebra*: Springer Science & Business Media, 2003.

[53] A. W. Leissa, "On a curve veering aberration," *Zeitschrift für angewandte Mathematik und Physik ZAMP,* vol. 25, pp. 99-111, 1974.

[54] A. Gallina, L. Pichler, and T. Uhl, "Enhanced meta-modelling technique for analysis of mode crossing, mode veering and mode coalescence in structural dynamics," *Mechanical systems and signal processing,* vol. 25, pp. 2297-2312, 2011.

[55] T. Igusa, "Critical Configurations Of Systems Subjected To Wide-Band Input," *Journal of Sound and Vibration,* vol. 168, pp. 525-541, 1993.

[56] C. Pierre, "Mode localization and eigenvalue loci veering phenomena in disordered structures," *Journal of Sound and Vibration,* vol. 126, pp. 485-502, 1988.

[57] L. Meirovitch, "Analytical methods in vibration," *New York, NY.: The Mcmillan Company,* vol. 19, 1967.

[58] D. K. Miu, *Mechatronics*: Springer, 1993.

[59] A. M. Rankers, *Machine Dynamics in Mechatronic Systems: An Engineering Approach,* 1997.

[60] E. Coelingh, T. J. de Vries, and R. Koster, "Assessment of mechatronic system performance at an early design stage," *Mechatronics, IEEE/ASME Transactions on,* vol. 7, pp. 269-279, 2002.

[61] M. Tohyama, "Room Transfer Function," in *Sound and Signals*, ed: Springer, 2011, pp. 243-304.

[62] P. Duffour and J. Woodhouse, "Instability of systems with a frictional point contact. Part 1: basic modelling," *Journal of Sound and Vibration,* vol. 271, pp. 365-390, 2004.

[63] D. K. Miu, "Physical Interpretation of Transfer Function Zeros for Simple Control Systems With Mechanical Flexibilities," *Journal of Dynamic Systems, Measurement, and Control,* vol. 113, pp. 419-424, 1991.

[64] J. Chandrasekar, J. B. Hoagg, and D. S. Bernstein, "On the zeros of asymptotically stable serially connected structures," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, 2004, pp. 2638-2643.

[65] H. J. Van de Straete, "Physical meaning of zeros and transmission zeros from bond graph models," Massachusetts Institute of Technology, 1995.

[66] G. C. Calafiore, "A subsystems characterization of the zero modes for flexible mechanical structures," in *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, 1997, pp. 1375-1380.

[67] B. Wie and A. E. Bryson Jr, "Modeling and control of flexible space structures," 1981.

[68] Y. J. Lee and J. L. Speyer, "Zero locus of a beam with varying sensor and actuator locations," *Journal of Guidance, Control, and Dynamics*, vol. 16, pp. 21-25, 1993.

[69] S. S. Aphale, A. J. Fleming, and S. R. Moheimani, "Integral resonant control of collocated smart structures," *Smart Materials and Structures*, vol. 16, p. 439, 2007.

[70] M. Vakil, R. Fotouhi, and P. Nikiforuk, "On the zeros of the transfer function of flexible link manipulators and their non-minimum phase behaviour," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 224, pp. 2083-2096, 2010.

[71] D. F. Enns, "Rocket stabilization as a structured singular value synthesis design example," *IEEE Control Systems*, vol. 11, pp. 67-73, 1991.

[72] L. Cui, C. Okwudire, and S. Awtar, "Complex Non-Minimum Phase Zeros in the Dynamics of Double Parallelogram Flexure Module Based Flexure Mechanisms," in *ASME 2016 Dynamic Systems and Control Conference*, Minneapolis, Minnesota, 2016.

[73] J. B. Hoagg and D. S. Bernstein, "Nonminimum-phase zeros: Much to do about nothing - Classical control revisited. Part II," *IEEE Control Systems Magazine*, vol. 27, pp. 45-57, 2007.

[74] M. Vidyasagar, "On undershoot and nonminimum phase zeros," *IEEE Transactions on Automatic Control*, vol. 31, pp. 440-440, 1986.

[75] B. Delabarra, "On undershoot in SISO systems," *IEEE Transactions on Automatic Control*, vol. 39, pp. 578-581, 1994.

[76] T. Damm and L. N. Muhirwa, "Zero crossings, overshoot and initial undershoot in the step and impulse responses of linear systems," *IEEE Transactions on Automatic Control*, vol. 59, pp. 1925-1929, 2014.

[77] J. Stewart and D. E. Davison, "On overshoot and nonminimum phase zeros," *IEEE Transactions on Automatic Control*, vol. 51, pp. 1378-1382, 2006.

[78] A. Preumont, *Vibration control of active structures: an introduction* vol. 179: Springer Science & Business Media, 2011.

[79] N. S. Nise, *CONTROL SYSTEMS ENGINEERING, (With CD)*: John Wiley & Sons, 2007.

[80] L. Cui, Y. Zhu, K. Yang, X. Li, and D. Yu, "Modeling and analyzing the main modes of an ultra-precision stage with lumped parameter method," in *Modelling, Identification & Control (ICMIC), 2012 Proceedings of International Conference on*, 2012, pp. 16-21.

[81] H. Butler, "Position control in lithographic equipment," *IEEE Control Systems Magazine*, vol. 31, pp. 28-47, 2011.

[82] B. Mace and E. Manconi, "Mode veering in weakly coupled systems," in *Proceedings of the ISMA conference*, 2012.

[83] X. L. Liu, "BEHAVIOR OF DERIVATIVES OF EIGENVALUES AND EIGENVECTORS IN CURVE VEERING AND MODE LOCALIZATION AND THEIR RELATION TO CLOSE EIGENVALUES," *Journal of Sound and Vibration*, vol. 256, pp. 551-564, 2002.

[84] T. Igusa, J. Achenbach, and K.-W. Min, "Resonance characteristics of connected subsystems: General configurations," *Journal of Sound and Vibration*, vol. 146, pp. 423-437, 1991.

[85] K. M. Xu and T. Igusa, "Dynamic Characteristics of Multiple Substructures with Closely Spaced Frequencies," *Earthquake Engineering & Structural Dynamics*, vol. 21, pp. 1059-1070, Dec 1992.

[86] P.-T. Chen and J. H. Ginsberg, "On the Relationship Between Veering of Eigenvalue Loci and Parameter Sensitivity of Eigenfunctions," *Journal of Vibration and Acoustics*, vol. 114, pp. 141-148, 1992.

[87] C. Pierre and P. D. Cha, "Strong Mode Localization in Nearly Periodic Disordered Structures," *AIAA journal*, vol. 27, pp. 227-241, 1989.

[88] C. Pierre and E. H. Dowell, "Localization of vibrations by structural irregularity," *Journal of Sound and Vibration*, vol. 114, pp. 549-564, 1987.

[89] A. Chandrashaker and S. Adhikari, "Vibration Localization of Rotationally Periodic Structures," in *Vibration Engineering and Technology of Machinery*, ed: Springer, 2015, pp. 865-878.

[90] M. P. Castanier and C. Pierre, "Modeling and analysis of mistuned bladed disk vibration: current status and emerging directions," *Journal of Propulsion and Power*, vol. 22, pp. 384-396, 2006.

[91] M. Nikolic, *New insights into the blade mistuning problem*: University of London, 2007.

[92] Y. Chen and I. Shen, "Mathematical insights into linear mode localization in nearly cyclic symmetric rotors with mistune," *Journal of Vibration and Acoustics*, vol. 137, p. 041007, 2015.

[93] X. Chen and A. Kareem, "Curve veering of eigenvalue loci of bridges with aeroelastic effects," *Journal of Engineering Mechanics*, vol. 129, pp. 146-159, 2003.

[94] C. E. Okwudire and J. Lee, "Minimization of the residual vibrations of ultra-precision manufacturing machines via optimal placement of vibration isolators," *Precision engineering*, vol. 37, pp. 425-432, 2013.

[95] J. L. du Bois, N. A. Lieven, and S. Adhikari, "Localisation and curve veering: a different perspective on modal interactions," in *27th Conference and Exposition on. Structural Dynamics (IMAC XXVII), Orlando, FL, Feb*, 2009, pp. 9-12.

[96] J. L. Du Bois, S. Adhikari, and N. A. Lieven, "On the quantification of eigenvalue curve veering: a veering index," *Journal of Applied Mechanics*, vol. 78, p. 041007, 2011.

[97] C. Acar and A. Shkel, *MEMS vibratory gyroscopes: structural approaches to improve robustness*: Springer Science & Business Media, 2008.

[98] Wikipedia. *Torsion constant*. Available: https://en.wikipedia.org/wiki/Torsion_constant

[99] D. B. Hiemstra, G. Parmar, and S. Awtar, "Performance tradeoffs posed by moving magnet actuators in flexure-based nanopositioning," *Mechatronics, IEEE/ASME Transactions on*, vol. 19, pp. 201-212, 2014.

[100] S. Awtar and D. D. Mariappan, "Experimental measurement of the bearing characteristics of straight-line flexure mechanisms," *Precision engineering*, vol. 49, pp. 1-14, 2017.

[101] X. Zhang, B. Koo, S. M. Salapaka, J. Dong, and P. M. Ferreira, "Robust control of a MEMS probing device," *IEEE/ASME Transactions on Mechatronics*, vol. 19, pp. 100-108, 2014.

[102] L. Cui, D. Mariappan, and S. Awtar, "Experimental Validation of Complex Non-Minimum Phase Zeros in Flexure Mechanism Dynamics," in *American Society of Precision Engineering, 31st Annual Meeting*, Portland, Oregon, US, 2016.

[103] R. M. Panas and J. B. Hopkins, "Eliminating Underconstraint in Double Parallelogram Flexure Mechanisms," *Journal of mechanical Design*, vol. 137, p. 092301, 2015.

[104] N. K. Roy and M. A. Cullinan, "Design of a Flexure Based XY Precision Nanopositioner with a Two Inch Travel Range for Micro-Scale Selective Laser Sintering," in *American Society of Precision Engineering, 31st Annual Meeting*, Portland, Oregon, US, 2016.

[105] K. K. Varanasi, "ibration damping using low-wave-speed media with applications to precision machines," Doctoral dissertation, Massachusetts Institute of Technology, 2004.

[106] Y. Yang and M. A. Butt, "Vibration Attenuation by Eddy Current with Different Combinations of Magnets," in *American Society of Precision Engineering, 31st Annual Meeting*, Portland, Oregon, US, 2016.